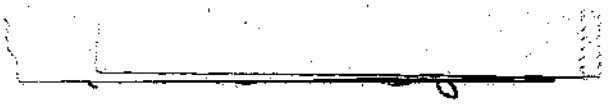
In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.



7/25/68

# A DISCRETE NONLINEAR RECURSIVE ESTIMATOR

A THESIS

Presented to

The Faculty of the Division of Graduate

Studies and Research

By

Richard E. Bryan

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

in the School of Electrical Engineering

Georgia Institute of Technology

June, 1972

A DISCRETE NONLINEAR RECURSIVE ESTIMATOR

Approved:

Roger P. Webb, Chairman

James E. Brown, III

Aubrey M. Bush

Date approved by Chairman: May 31, 1972

ACKNOWLEDGEMENTS

First, I would like to thank Mrs. Jean Shankel for the faithful rendering of her good services in typing this manuscript.

Secondly, I wish to thank the members of my reading committee, Dr. James E. Brown, III, Dr. Aubry M. Bush, and Dr. James R. Rowland, whose thorough critiques aided greatly in polishing this work. Also, I wish to acknowledge the encouragement given me by the personal friendship of my colleagues, Dr. John Scardina and Dr. Robert Howerton.

Most important of all, I offer the sincerest appreciation to my advisor, Dr. Roger P. Webb, who offered the full measure of his insight and understanding in guiding me through this work.

Finally, I offer lasting gratitude for the quiet but forceful patience of Bunny J. Flowers.

TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF ILLUSTRATIONS

Figure Page

## SUMMARY

The objective of this dissertation is to develop an estimator algorithm for certain estimation problems with an incomplete problem model which may contain nonlinearities and nongaussian statistics. To this end a problem model is completely defined using the given assumptions and additional assumed constraints, an estimator algorithm is derived based on this model, and the suboptimal nature of the resulting estimator is analyzed to provide additional practical results.

The only initial assumptions are that the system dynamics are known and that the control input and the corrupting noise are bounded with known bounds. In order to derive an estimator based on these assumptions, it is necessary to define a complete model of the problem. This model is constructed by assuming additional constraints on the problem. These constraints are that the signal and corrupting noise are uniformly distributed and uncorrelated and that the corrupting noise and control input are white. This model allows optimal estimation in cases where all constraints are met and adequate suboptimal estimation in many other cases. For instance, in many cases where the signal and corrupting noise statistics are bounded but unknown, a uniform distribution assumption yields better suboptimal results than a linear estimator based on the assumption of Gaussian statistics.

The optimal estimator algorithm derivation based on the completed problem model follows directly from the well-known principle that the mean of the signal density conditioned on the measurement yields the

minimum mean square error (MMSE) estimate. In deriving the optimal estimator algorithm it is necessary to consider the case of the linear plant and the case of the nonlinear plant separately although both are derived under identical constraints. The separate derivations are necessitated by computational problems arising from the nonlinear plant. In both cases the resulting algorithm is nonlinear.

Suboptimal experimental work compares the resulting nonlinear estimator to the well-known linear Kalman estimator by establishing a set of conditions suboptimal to both estimators and then varying major controllable parameters in the measurement generating system to determine conditions under which the nonlinear estimator performance exceeds that of the Kalman estimator.

# CHAPTER I

# INTRODUCTION

The purpose of this research is to develop new techniques for estimating discrete signals when the only measurement available is corrupted by additive noise. In particular, a minimum mean square error (MMSE) estimator algorithm is derived for a class of problems with nonlinearities and nongaussian statistics. The resulting algorithm is nonlinear.

The estimator derived here is primarily intended to be used in certain cases where the signal and corrupting noise statistics are unknown. For instance, in cases where the system input and corrupting noise are bounded with known bounds, the estimator can be implemented using only the known bounds and assuming a uniform distribution for lack of better information. It is also possible that the estimator can be used optimally in a limited number of cases.

## Statement of the Problem

The estimation problem considered in this dissertation consists of three parts:

1. Definition of a complete set of problem constraints which do not exclude nonlinearities or nongaussian statistics.

2. Derivation of a MMSE algorithm based on the constraints in 1.

3. Experimentation with suboptimal problems to determine further practical applications.

A model for the problem is shown in Figure 1. Estimates, $\hat{x}$, are made from measurements, y, which are corrupted by additive noise, v. The signal to be estimated, x, and the corrupting noise may be continuous or discrete; but the measurements will be discrete; and the resulting estimator will be defined by a nonlinear recursive estimation rule. It will be assumed that the signal generating system dynamics are known and that x and v are bounded and statistically independent.

The most commonly used estimator is the linear Kalman estimator; however, this estimator is designed primarily for problems with Gaussian input and corrupting noise and linear signal generating system. For these assumptions, the linear Kalman estimator is optimum over all estimators, linear or nonlinear, for all convex cost functions.

Many practical problems arise involving nonlinearities, nongaussian statistics or even unknown statistics. It is frequently possible to linearly approximate the system, assume gaussian statistics, and achieve adequate suboptimal results using a linear estimator. The work in this dissertation describes an estimator which provides optimal estimation for certain of these problems and another suboptimal alternative for many other cases.

## History of the Problem

The general estimation problem is that of estimating a signal when the only measurement available is corrupted by additive noise. Several different solutions to the estimation problem are available, each being based on different sets of assumed constraints on the nature of the signal to be estimated and the corrupting noise.

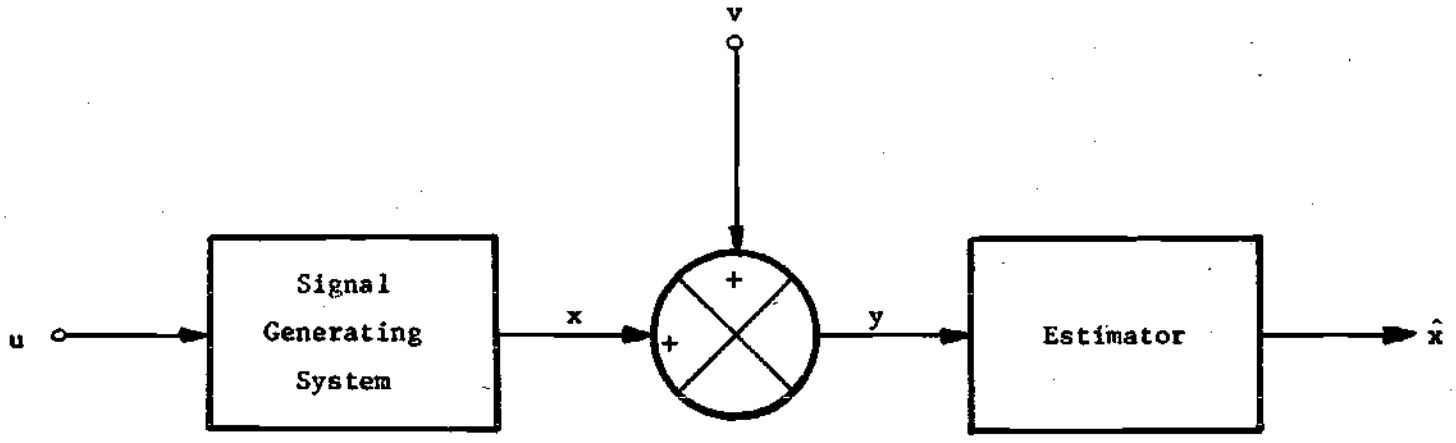



Figure 1. Estimator Problem Model

The most commonly known solutions are the linear estimators of Weiner [1] and Kalman [2]. Since the initial assumptions for both estimators in their basic form are equivalent and since both estimators minimize mean square error between the actual and estimated signals, the results are equivalent in steady state.

The major difference in the two estimators is that the Weiner estimator, derived in the frequency domain, is a steady state estimator, i.e., initialized at $t = -\infty$; and the Kalman estimator derived in state space, includes transient effects, i.e., initialized at $t = 0$. For the case where the signal is generated by a stationary system, the Kalman estimator converges to the Weiner estimator. Booton [3] considers the applicability of the Weiner method to the nonstationary case; however, the results are for the most part not useful in a practical sense. The Kalman estimator handles the nonstationary case easily with no additional complications.

Basically the Weiner-Kalman solution is predicated on the assumptions that the signal to be estimated is the output of a linear system driven by Gaussian white noise and the additive corrupting noise is also Gaussian and white. This particular set of assumptions gives the most basic solution. Various extensions, such as the use of colored input noise and the use of input and corrupting noise are commonly found in the literature. Shaw [4] considers Kalman extensions, and Middleton [5] considers Weiner extensions. In all of these extensions the Gaussian constraint on input and corrupting noise and the linearity constraint on the signal generating system are maintained.

Most of the post Kalman work has proceeded in the following areas:

1. Applications of the Kalman estimator such as the work of Fagin [6] with inertial guidance, the work of Schmidt [7] with non-evasive tracking and the work of Park [8] with closed loop control systems.

2. Analysis of the sensitivity of the Kalman estimator to errors in the assumed data such as the work of Heffes [9], Nishimura [10], and Fitzgerald [11].

3. Attempts to establish more general estimation techniques such as the Bayesian approach of Lee [12]. It should be noted the Weiner-Kalman estimator is more rigidly constrained than the Bayesian estimator although Lee shows that the Bayesian estimator assumes a Kalman form when the Kalman constraints are applied.

4. Derivation of estimators for problems with a nonlinear system function. This group of estimators may be dichotomized by considering those which resort to linearization such as Cox [13] and Athans, et al. [14] and those which do not use linearization such as Frost, et al. [15]. Estimators in the first category generally use the concept of the "extended Kalman estimator" which involves linearizing the system function about the most recent estimate. Several estimation techniques for problems with nonlinearities are compared by Mehra [16].

5. Solutions for particular types of problems in which the given information conflicts with the Kalman constraints or is insufficient such as the work of Kuo [17] and Schweppe [18]. The works of Kuo and Schweppe are important since they have both attacked the problem of estimation with undefined statistical parameters.

The philosophy behind the work in this fifth area is also important. The Gaussian restriction of the Kalman estimator precipitates the further restriction of the linear signal generating system. Together these constraints preclude problems with bounded variables, unknown statistics, or nonlinearities. Many practical problems do not fit such rigid constraints, and this has promoted the design of estimators for particular types of problems which do not make use of the Kalman assumptions.

For instance, Schweppe, motivated by a desire to track vehicles attempting evasive maneuvers, has derived an estimator without resorting to any statistics whatsoever. His initial assumptions are that the vehicle input and corrupting noise are bounded with known bounds and that the vehicle dynamics are completely known. He then proceeds to derive an estimator algorithm using the idea that the estimate lies in the intersection of the reachable states and the states which could have given the observed measurement. Because the estimator is not derived with regard to any optimizing criterion and because of the nature of the resulting estimator, it is impossible to analyze its effectiveness in reducing error except experimentally.

Starting with essentially the same assumptions as Schweppe, Kuo has developed an adaptive estimator which learns the moments of the signal and corrupting noise. The resulting estimator does not include transient effects and is computationally cumbersome to the extent that it cannot be used for on-line work.

The work in this dissertation also begins with the same assumptions as those used by Schweppe. Rather than proceeding with a nonsta-

tistical estimator, a complete statistical model is then defined using the most logical assumptions in lieu of additional given information; and an estimator algorithm is derived. The intersection idea of Schweppe is used to the extent that it leads to a conditional mean (Bayesian) estimator when probability structure is added to the signal and corrupting noise definitions.

# CHAPTER II

## OPTIMAL ESTIMATOR FOR LINEAR PLANT

The general estimation problem is so formidable that it may never be solved in its most basic form. Most work in the estimation area has proceeded along the lines of first defining a set of constraints which isolates a subset of the set of general estimation problems and then proceeding to derive a MMSE estimator for the subset. This dissertation follows such a format and in this chapter the constraints are defined and a MMSE estimator algorithm is derived for systems with linear plants. A similar algorithm is derived for certain systems with nonlinear plants in the following chapter. In both cases the initial constraints are the same and the resulting algorithm is nonlinear. It is necessary to consider the case of a nonlinear plant separately not because it violates any of the constraints used in the case of the linear plant but because of computational problems arising from the iteration of the system bounds from one time instant to the next when the plant is nonlinear.

### Constraints

The general constraints assumed at the initiation of the problem are that the signal generating system dynamics are known and that the control input, u, and the corrupting noise, v, are bounded with known bounds. These conditions are essentially the same as those used by Schweppe [18]. At this point the work of Schweppe and the work in this dissertation diverge. Schweppe proceeds to derive a nonstatistical estimator without

any additional knowledge of the problem. The procedure used here is to completely define a statistical model of the problem by assuming additional constraints on the system and then to derive a MMSE estimator algorithm.

The initial assumptions defining the basic problem are quite simple and totally nonstatistical in nature. The additional assumptions, which give probability structure to the problem, are added in order to effect an estimator algorithm based on the Bayesian principle, i.e., the principle that the MMSE estimate is the mean of the signal density conditioned on the measurement. These additional assumed properties are chosen to be logically consistent with practical problems and so as not to exclude nongaussian or nonlinear problems.

In order to completely define the problem model, the following assumed properties are added to the given characteristics:

1. The additive corrupting noise is uniformly distributed zero mean white noise.

2. The control input forcing the signal generating system is zero mean white noise.

3. The signal to be estimated is a bounded uniformly distributed stochastic process.

4. The signal and corrupting noise are statistically independent.

For the most part these additional constraints are straightforward and classical estimation problem constraints. For instance, the assumptions that the corrupting noise and control input are zero mean white noise and that the signal and corrupting noise are statistically independent are basic and not unduly restrictive. Also, that the signal is

bounded follows from the given bounded input constraint and the stability of the signal generating system.

The assumption that the signal and corrupting noise are uniformly distributed follows from the fact that these quantities are bounded with no other information available regarding their statistical distribution. This result is intuitively satisfactory but can also be concluded using entropy considerations. See [25], p. 334.

The total set of resulting constraints are the following:

1. The signal generating system dynamics are known.

2. The additive corrupting signal is bounded zero mean white noise uniformly distributed with known bounds.

3. The control input forcing the signal generating system is bounded zero mean white noise with known bounds.

4. The signal to be estimated is a bounded uniformly distributed stochastic process.

5. The signal and corrupting noise are statistically independent.

The basic estimator algorithm is derived on the basis of these constraints. Some constraints may be partially relaxed for certain special cases. In these cases the algorithm is modified so that the resulting estimate remains optimal. One such case is considered at the end of this chapter.

It should be noted that the total set of constraints is extremely rigid and there are few practical cases where the resulting estimator could be used optimally; however, there are many practical cases where the estimator could be used suboptimally to great advantage. For instance, many practical problems involve bounded signal and corrupting noise with

little information available concering the statistics of the signal and the noise. For such cases it may be assumed that the signal and corrupting noise are uniformly distributed. Unlikely as it may be that the result will be optimal, the suboptimal performance will frequently exceed that of a linear estimator used in the same application. An attempt is made to be more definitive with regard to suboptimal use of the estimator algorithm in Chapter III.

## Estimator Design

Given the preceeding constraints it is necessary to devise an estimator algorithm which gives a MMSE estimate. To insure that the estimate minimizes mean square error (MSE) the algorithm will choose as an estimate the mean of the signal, x, conditioned on the measurement, y. For the constraints chosen, the density of x conditioned on y is symmetrical and bounded.* Therefore, the conditioned mean of x given y is simply the geometric center of the nonzero probability region of x given y and can be calculated knowing only the bounds, c(i) and d(i), or endpoints of the region. The procedure to be incorporated into an algorithm is to calculate at each time instant the exact bounds of x given y.

The sequence of operations performed by the algorithm is outlined in the following:

1. Given the optimal estimate $\hat{x}(i)$, and the lower and upper bounds, c(i) and d(i) respectively at some time instant, say $t = t_i$, the set of reachable states at $t = t_{i+1}$ are found in terms of bounds for the system states. These bounds are called initial or maximum bounds.

*See Appendix A.

2. Given the measurement, y(i+1), the set of reachable states is reduced to the set which could be reached and also give the measurement taken. In conjunction with this concept, the bounds are separated into three categories: those which could not possibly be affected by the measurement (bounds for unmeasured variables), those which could be affected by the measurement but are not because of the particular value of the measurement, and those which are affected by the measurement.

3. Once the proper adjustments are made to the bounds which are affected, the three categories of bounds are combined to give the set of final corrected bounds for the system states at $t = t_{i+1}$, and the estimate is then based on these refined bounds.

Such an algorithm is defined in the following for the general n-dimensional problem in vector-matrix notation.

At some time instant, $t_i$, the lower and upper bounds on x(i) given y(i) are respectively c(i) and d(i). The system dynamics and measurement system are defined respectively by

$$x(i+1) = F\,x(i) + G\,u(i+1) \tag{2.1}$$

and

$$y(i+1) = H\,x(i+1) + v(i+1) \tag{2.2}$$

where x(i) is nxl, F is nXn, G is nxp, u(i) is pxl, H is nxn, v(i) is nxl, and y(i) is nxl for all i. There are m measured states, $m \leq n$, and these states are measured individually rather than in linear combinations. Therefore, the H matrix has a "1" on the main diagonal for

each measured state and zeros everywhere else, and both y and v have n-m zero elements corresponding to the n-m unmeasured states. Also note that the m nonzero states of v are bounded symmetrically about zero by the vector b on the upper side and the vector -b on the lower side, and that b has n-m zero elements corresponding to the n-m unmeasured states.

Now define an nxn matrix $F^+$ which is simply the transition matrix, F, with all negative elements replaced by zero and an nxn matrix $F^-$ which is the transition matrix with all positive elements replaced by zero. Also define a quantity $\Delta(i)$ given by

$$\Delta(i) = d(i) - \hat{x}(i) \tag{2.3}$$

where $\hat{x}(i)$ is the optimal estimate at $t_i$. Since $\hat{x}(i)$ is the geometric center of the bounds at $t_i$,

$$-\Delta(i) = c(i) - \hat{x}(i). \tag{2.4}$$

In the algorithm it is first necessary to determine the absolute upper and lower limits which could be reached by x at $t_{i+1}$ given the limits c(i) and d(i) at $t_i$. As will be subsequently shown, these reachable limits at $t_{i+1}$ are given by

$$c'(i+1) = F\,\hat{x}(i) - F^+\,\Delta(i) + F^-\,\Delta(i) + G\,\underline{u} \tag{2.5}$$

and

$$d'(i+1) = F\,\hat{x}(i) + F^+\,\Delta(i) - F^-\,\Delta(i) + G\,\bar{u} \tag{2.6}$$

where $\underline{u}$ and $\bar{u}$ are respectively the lower and upper limit vectors for

the system input, u. Note that $\underline{u} = -\bar{u}$, i.e., the limit vectors are symmetrical about zero and that the individual elements of these limit vectors may vary with time as long as they remain symmetrical about zero and are known at each time instant.

To provide some insight into equations (2.6) and (2.7), consider the following: At $t_i$ each individual element of x(i) has a lower and upper bound, and some elements may lie near their lower bound whereas others may lie near their upper bound. In order to find the absolute minimum reachable bound at $t_{i+1}$, c'(i+1), it is necessary to assume that each element in x(i) lies in the position between its lower and upper bound which will give the maximum negative contribution to each element in c'(i+1). This cannot be accomplished by simply iterating c(i) through the transition matrix and using $c'(i+1) = F\ c(i) + G\ \underline{u}$. Consider a particular element, say $c'_j(i+1)$, and suppose the $h^{th}$ element in $j^{th}$ row of F is negative, i.e., $f_{jh} < 0$. In order to give the maximum negative contribution to $c'_j(i+1)$ it is necessary to assume that $x_h(i)$ lies at its upper bound, $d_h(i)$. All of this can be accomplished for every term in every element of c'(i+1) by setting

$$c'(i+1) = F^{+}\ c(i) + F^{-}\ d(i) + G\ \underline{u}. \tag{2.7}$$

Similarly the absolute upper limit can be written

$$d'(i+1) = F^{+}\ d(i) + F^{-}\ c(i) + G\ \underline{u}. \tag{2.8}$$

Substituting $c(i) = \hat{x}(i) - \Delta(i)$ and $d(i) = \hat{x}(i) + \Delta(i)$ into (2.7) and (2.8) and rearranging produces equations (2.5) and (2.6). Equations

(2.5) and (2.6) are used in the remaining derivation so that the final algorithm in this chapter may be conveniently compared to the final algorithm in the next chapter.

Before considering the effect of the measurement, y(i+1), on the bounds, it is necessary to isolate those bounds which can be affected by the measurement. This is accomplished using the measurement matrix H. Those bounds which can be affected by the measurement are given by

$$\tilde{c}(i+1) = H\, c'(i+1) \tag{2.9}$$

and

$$\tilde{d}(i+1) = H\, d'(i+1). \tag{2.10}$$

Then those bounds which cannot be affected by the measurement are given by

$$c^1(i+1) = (I-H)\, c'(i+1) \tag{2.11}$$

and

$$d^1(i+1) = (I-H)\, d'(i+1). \tag{2.12}$$

The effect of the measurement on the boundaries of states which can be affected by the measurement is shown pictorially in Figure 2. At $t_{i+1}$, the region reachable by $x_j(i+1)$ is bounded by $\tilde{c}_j(i+1)$ and and $\tilde{d}_j(i+1)$. Since $v_j(i+1)$ has an upper bound of $b_j$ and a lower bound $-b_j$, the measurement has lower and upper bounds respectively $\tilde{c}_j(i+1)$ $-b_j$ and $\tilde{d}_j(i+1) + b_j$. Suppose the measurement falls in the region between $\tilde{c}_j(i+1) + b_j$ and $\tilde{d}_j(i+1) + b_j$ such as shown by $y_j(i+1) = k$. The

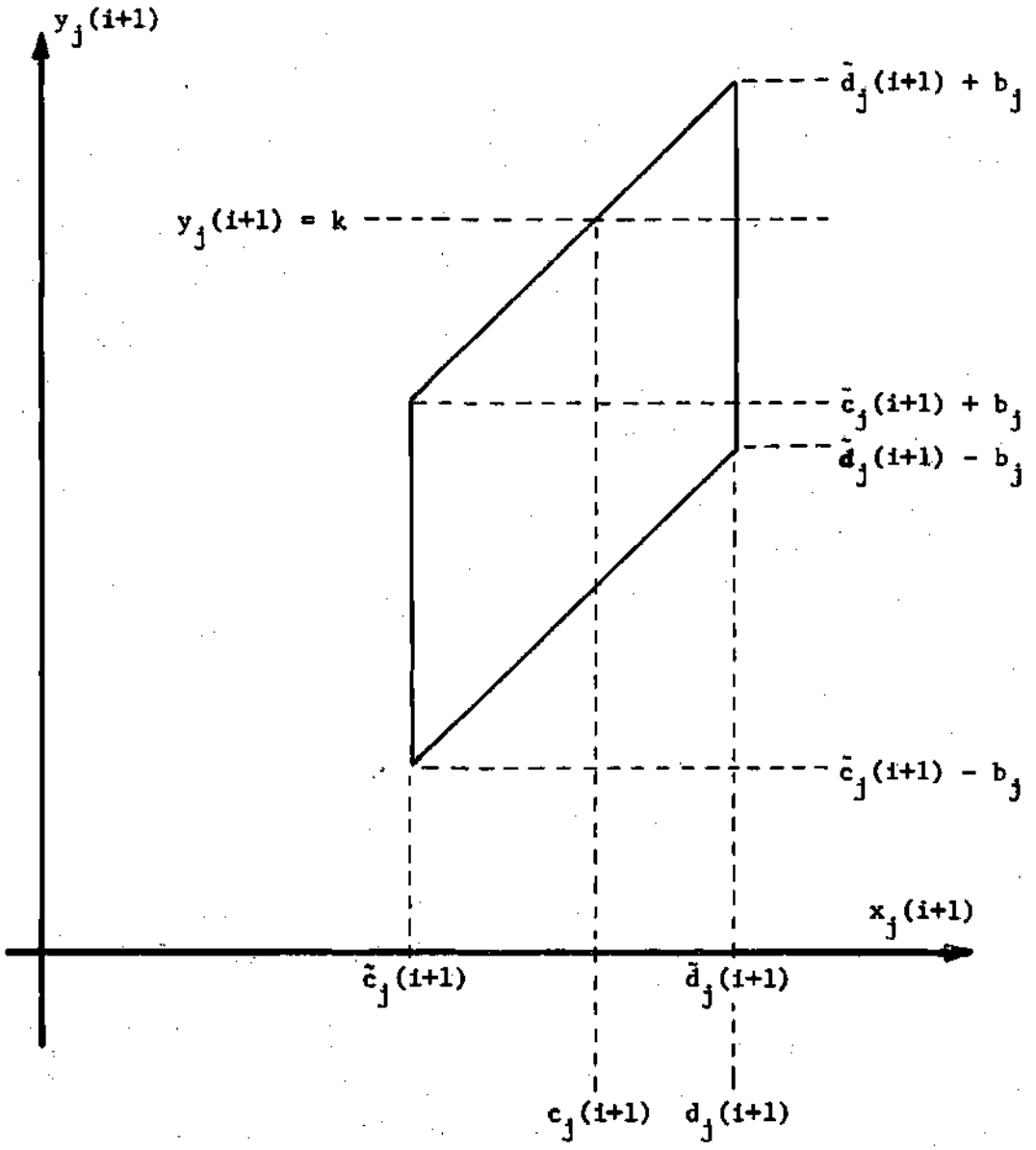



Figure 2. Effect of Measurement on Reachable Set Bounds

system states for $x_j(i+1)$ which could have yielded such a measurement have lower and upper bounds respectively $y_j(i+1) - b_j$ and $y_j(i+1) + b$. Taking the intersection of the reachable states and those states which could have yielded the measurement as the region containing $x_j(i+1)$ given $y_j(i+1)$ it is found that the bounds for this region are

$$c_j(i+1) = y(i+1) - b_j \tag{2.13}$$

and

$$d_j(i+1) = \tilde{d}_j(i+1). \tag{2.14}$$

In other words the measurement has no effect on the upper bound for $x_j(i+1)$ but improves the lower bound. Similarly a measurement falling between $\tilde{d}_j(i+1) - b_j$ and $\tilde{c}(i+1) - b_j$would improve the upper bound but not the lower bound, and the final bounds would be given by

$$c_j(i+1) = \tilde{c}_j(i+1) \tag{2.15}$$

and

$$d_j(i+1) = y_j(i+1) + b_j. \tag{2.16}$$

Note that a measurement falling between $\tilde{d}(i+1) - b_j$ and $\tilde{c}_j(i+1) + b_j$ would not improve either bound. If, on the other hand, it should happen from the structure of the problem that $\tilde{d}(i+1) - b_j > \tilde{c}_j(i+1) + b_j$, then a measurement falling between $\tilde{c}(i+1) + b_j$ and $\tilde{d}(i+1) - b_j$ would improve both the upper bound and the lower bound.

Next in designing the estimator algorithm it is necessary to define decision matrices will be used to isolate those bounds which

are actually affected by the measurement. To this end define

$$\left.\begin{aligned} k_{cj}(i+1) &= 1,\ y_j(i+1) \leq \tilde{c}_j(i+1) + b_j \\ &= 0,\ \text{otherwise} \end{aligned}\right\} j\epsilon S$$
$$= 0,\ j\notin S \qquad (2.17)$$

and

$$\left.\begin{aligned} k_{dj}(i+1) &= 1,\ y_j(i+1) \geq \tilde{d}(i+1) - b_j \\ &= 0,\ \text{otherwise} \end{aligned}\right\} j\epsilon S$$
$$= 0,\ j\notin S \qquad (2.18)$$

where S is a set of m integers corresponding to the subscripts of the m measured states. These elements are used to form the main diagonals of the two decision matrices:

$$K_c(i+1) = \begin{bmatrix} k_{c_1}(i+1) & & 0 \\ & \ddots & \\ 0 & & k_{cn}(i+1) \end{bmatrix} \qquad (2.19)$$

and

$$K_d(i+1) = \begin{bmatrix} k_{d_1}(i+1) & & 0 \\ & \ddots & \\ 0 & & k_{dn}(i+1) \end{bmatrix} . \qquad (2.20)$$

The decision matrices have "1"'s on the main diagonal for bounds which could be affected by the measurement but are not because of the value of the measurement. These bounds may be written as

$$c^2(i+1) = K_c(i+1)\ \tilde{c}(i+1) \tag{2.21}$$

and

$$d^2(i+1) = K_d(i+1)\ \tilde{d}(i+1) \tag{2.22}$$

The bounds which can be affected by the measurement and are affected can be written, respectively,

$$\tilde{c}^3(i+1) = \{I-K_c(i+1)\}H\tilde{c}(i+1) \tag{2.23}$$

$$\tilde{d}^3(i+1) = \{I-K_d(i+1)\}H\tilde{d}(i+1). \tag{2.24}$$

Once the proper corrections are made to those bounds defined by (2.23) and (2.24), they may be written

$$c^3(i+1) = \{I-K_c(i+1)\}H\{y(i+1) - b\} \tag{2.25}$$

and

$$d^3(i+1) = \{I-K_d(i+1)\}H\{y(i+1) + b\}. \tag{2.26}$$

Now the three types of bounds may be combined in a single expression to give the final minimum bounds at $t_{i+1}$:

$$c(i+1) = c^1(i+1) + c^2(i+1) + c^3(i+1) \tag{2.27}$$

and

$$d(i+1) = d^1(i+1) + d^2(i+1) + d^3(i+1). \tag{2.28}$$

Making the correct substitutions in (2.27) and (2.28) yields

$$c(i+1) = (I-H)c'(i+1) + K_c(i+1)Hc'(i+1) + [I-K_c(i+1)]H[y(i+1)-b] \tag{2.29}$$

and

$$d(i+1) = (I-H)d'(i+1) + K_d(i+1)Hd'(i+1) + [I-K_d(i+1)]H[y(i+1)+b]. \tag{2.30}$$

Because of symmetry in the density of x conditioned on y, the MMSE estimate is

$$\hat{x}(i+1) = \frac{1}{2}\,[c(i+1) + d(i+1)]. \tag{2.31}$$

Substituting (2.29) and (2.30) into (2.31) yields

$$\begin{aligned}\hat{x}(i+1) = \frac{1}{2}\,\{&(I-H)c'(i+1) + K_c(i+1)Hc'(i+1)\\ &+[I-K_c(i+1)]H[y(i+1)-b]\\ &+ (I-H)d'(i+1)\\ &+ K_d(i+1)Hd'(i+1)\\ &+ [I-K_d(i+1)]H[y(i+1)+b]\}.\end{aligned} \tag{2.32}$$

Substituting (2.5) and (2.6) into (2.32) and simplifying gives the final MMSE estimator algorithm:

$$\hat{x}(i+1) = (I-H)\,F\hat{x}(i) + \frac{1}{2}\{K_d(i+1) + K_c(i+1)\}F\hat{x}(i)$$
$$+ \frac{1}{2}\{K_d(i+1) - K_c(i+1)\}\{F^{+}\Delta(i) + F^{-}\Delta(i) + G\bar{u}\}$$
$$+ \frac{1}{2}\{I-K_c(i+1)\}H\{y(i+1) - b\} \qquad (2.33)$$
$$+ \frac{1}{2}\{I-K_d(i+1)\}H\{y(i+1) + b\}.$$

A simplified block diagram of the estimator is shown in Figure 3. In this figure L'(i+1) represents the unreduced limits; $\tilde{L}(i+1)$ represents limits which can be reduced by the measurement; $\tilde{L}(i+1)$ represents limits which have been reduced by the measurement; and L(i+1) represents the final limits upon which the estimate is based. A more detailed block diagram is shown in Figure 4. An additional computed quantity is shown in this diagram, namely, E (i+1) which is an estimate of the MSE in $\hat{x}(i+1)$ at $t_{i+1}$. Since the density of x(i+1) conditioned on y(i+1) is uniform and since the expected value of the MSE in x(i+1) is the variance of x(i+1) conditioned on y(i+1)

$$E(i+1) = \frac{1}{12}\{d(i+1) - c(i+1)\}^2. \qquad (2.34)$$

## Extension

The constraints given earlier in this chapter and the algorithm derived on the basis of these constraints represent the estimator in its most basic form. It is possible to extend this algorithm to cover cases ruled out in the basic constraints. The extension to be covered in this section is the use of bandlimited system input noise.

Suppose the signal generating system is given as a transfer function $T_2(S)$ and suppose the system input, u, is derived by passing white

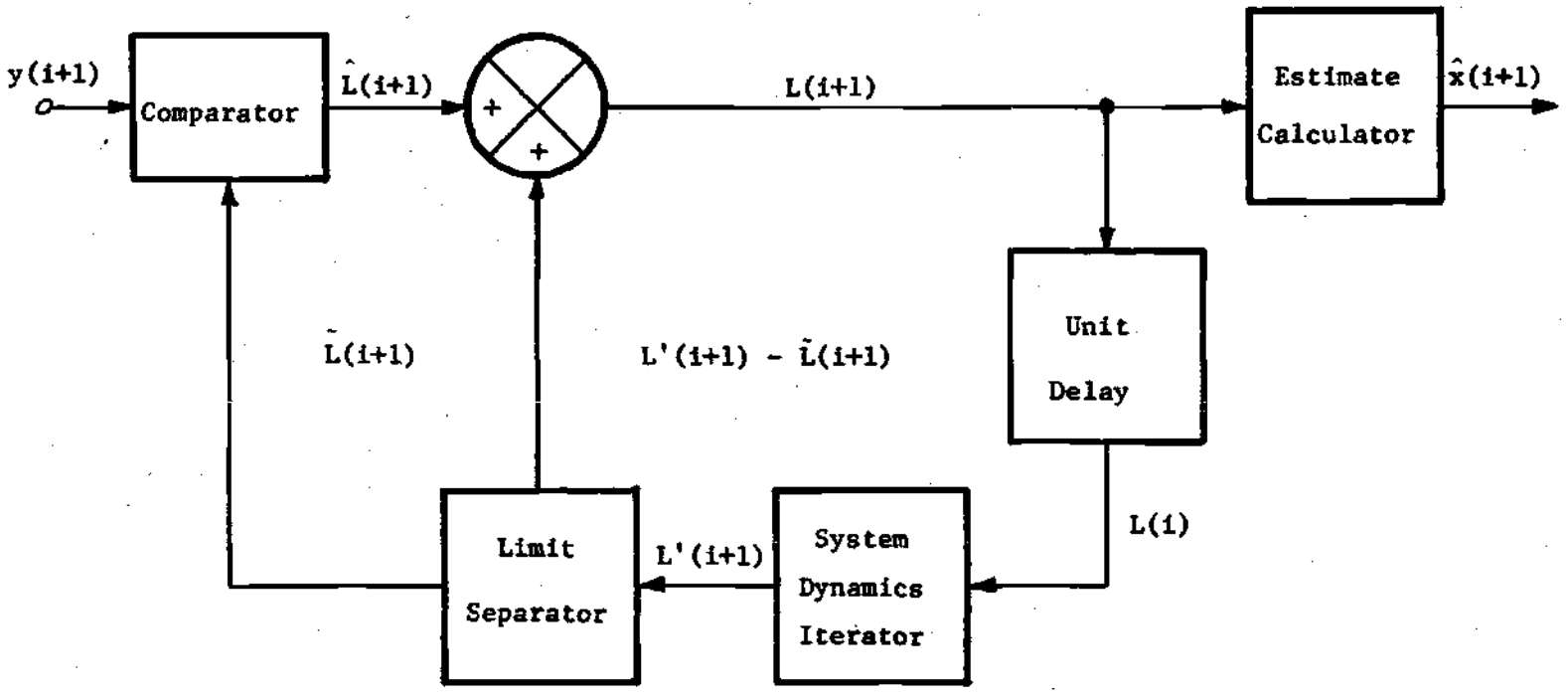



Figure 3. Estimator Simplified Block Diagram

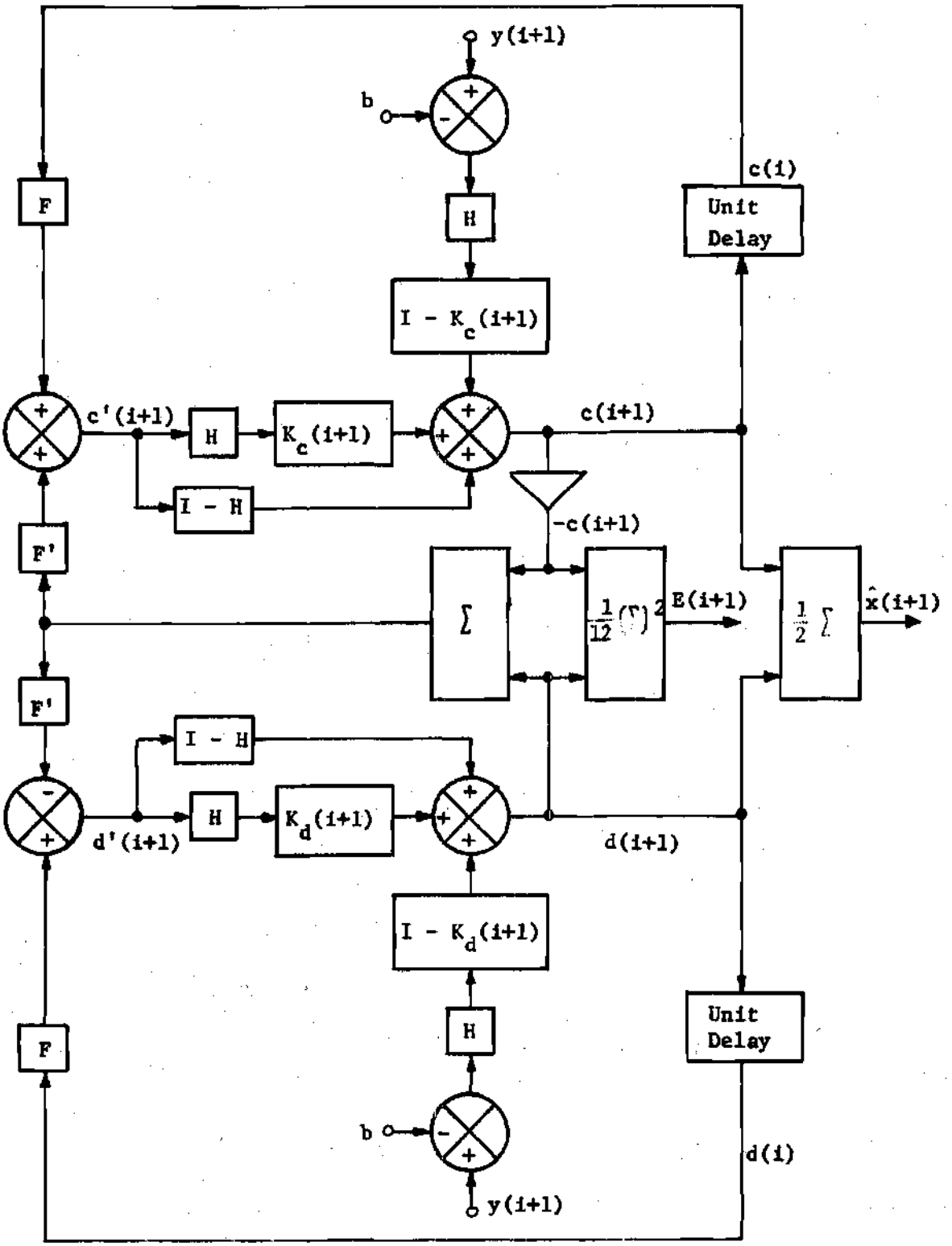



Figure 4. Estimator Block Diagram

noise, w, through a system with transfer function $T_1(S)$. Such a configuration is shown in Figure 5. To handle this problem the system input is considered to be w and the system transfer function is T(S) where

$$T(S) = T_1(S)\ T_2(S). \tag{2.35}$$

Now the basic algorithm can be used as given in (2.33) by transforming the given continuous system into a discrete system.

If the system with colored input noise is given in discrete form a different but analogous approach is used. For instance, suppose the signal generating system is given by

$$x(i+1) = F_2x(i) + G_2u(k) \tag{2.36}$$

and the input is generated by passing white noise, w(i), through the following system

$$u(i+1) = F_1u(i) + G_1w(i). \tag{2.37}$$

The two systems are combined using the following definitions

$$\underset{\sim}{x}(i+1) = \begin{bmatrix} x(i+1) \\ \hline u(i+1) \end{bmatrix} \tag{2.38}$$

$$F = \left[\begin{array}{c|c} F_1 & G_1 \\ \hline 0 & F_2 \end{array}\right] \tag{2.39}$$

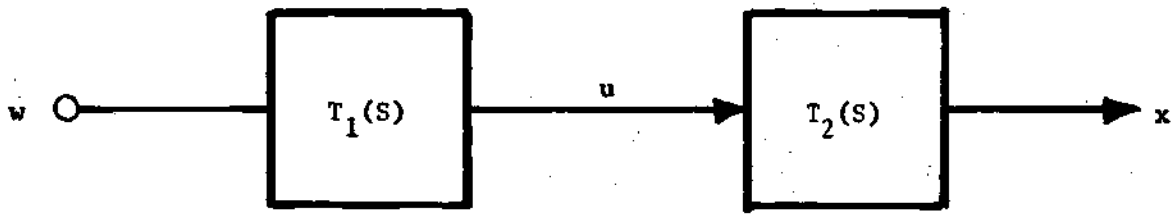



Figure 5. Model for Signal Generating System with Colored Noise Input

$$G = \begin{bmatrix} 0 \\ \hline G_2 \end{bmatrix} \tag{2.40}$$

$$\tilde{u}(i) = \begin{bmatrix} 0 \\ \hline w(i) \end{bmatrix} \tag{2.41}$$

Now the dynamics of the combined system are given by

$$\tilde{x}(i+1) = F\tilde{x}(i) + G\tilde{u}(i) \tag{2.42}$$

where $\tilde{x}(i+1)$ is the output vector and $\tilde{u}(i)$ is the white noise input vector and the basic algorithm can be used as given in (2.33).

# CHAPTER III

## OPTIMAL ESTIMATOR FOR NONLINEAR PLANT

In this chapter an estimator algorithm is derived for a system which is nonlinear in state, linear in control, and has a differentiable nonlinear system function. At each sample point the estimator uses linear pertubations about the previous estimate in order to arrive at a current estimate. The measured states are measured individually rather than in linear combinations as in the case of the linear plant.

The problem considered here is identical to the problem in Chapter II (including both structure and constraints) except that the given plant is represented by a nonlinear system function rather than a linear transition matrix. It is necessary to consider the system with nonlinear plant separately not because it violates the constraints used in the linear derivation but because of computational problems arising from the iteration of the system bounds from one time instant to the next. These problems are resolved by linearly iterating the bounds about a nominal trajectory with the final algorithm being very similar to that derived in Chapter II for the linear case.

The purpose in presenting the material in this chapter is simply to show that the estimator procedure will work for nonlinear systems, and no attempt is made to thoroughly examine the nonlinear case. The major points in this work are the derivation method used in this chapter and the previous chapter and the suboptimal examples shown in the next

chapter. It is hoped that future work will evolve an estimator algorithm for nonlinear plants without recourse to linearization. Such an algorithm is theoretically possible since nonlinear plants are not ruled out by the initial assumptions.

In the remainder of this chapter the optimal algorithm for the case of a nonlinear plant with linear control is derived, and this algorithm is compared to the algorithm for the case of a linear plant with linear control.

## Estimator Design

The sequence of operations performed by the algorithm is essentially the same as outlined at the beginning of Chapter II. Notation used for various types of system bounds is identical to that used in Chapter II and is briefly reviewed in the following: at some time, $t_i$, the final corrected lower and upper bounds are noted respectively by c(i) and d(i); the initial or "rough" bounds at $t_{i+1}$ are noted by c'(i+1) and d'(i+1); the initial bounds for unmeasured variables are noted by $c^1(i+1)$ and $d^1(i+1)$, and the bounds for measured states are noted by $\tilde{c}(i+1)$ and $\tilde{d}(i+1)$; bounds for measured states which are not affected by the particular measurement taken are noted by $c^2(i+1)$ and $d^2(i+1)$, and bounds for measured states which are affected by the particular measurement taken are noted by $\tilde{c}^3(i+1)$ and $\tilde{d}^3(i+1)$; and finally, the affected bounds are noted by $c^3(i+1)$ and $d^3(i+1)$ after the proper corrections have been made.

The given system has the following structure:

$$x(i+1) = f\{x(i)\} + Gu(i) \qquad (3.1)$$

$$y(i+1) = Hx(i+1) + v(i+1). \tag{3.2}$$

The system output, x(i), is n x 1; the system input, u(i), is p x 1; the nonlinear system function, f(·), is n x 1; G is n x p; the measurement, y(i), is n x 1; and the corrupting noise, v(i), is n x 1. There are m measured states, $m \leq n$, and these states are measured individually rather than in linear combinations. Therefore, the H matrix is n x n with a "1" on the main diagonal for each measured state and zeroes everywhere else, and both y and v have n - m zero elements corresponding to the n - m unmeasured states.

Before proceeding with the estimator algorithm derivation it is necessary to define certain elements which will be used in this derivation:

Define A(i) to be the n x n Jacobian matrix of $f\{x(i)\}$ at $t = t_i$ with elements

$$\{A(i)\}_{\alpha\beta} = \left.\frac{\partial f_\alpha\{x(i)\}}{\partial x_\beta}\right|_{x(i) = \hat{x}(i)} \tag{3.3}$$

where

$$\alpha,\beta = 1, 2, \ldots, n.$$

Define $A^+(i)$ to be an n x n matrix containing only the positive elements of A(i) with all other elements replaced by zero. Define $A^-(i)$ to be an n x n matrix containing only the negative elements of A(i) with all other elements replaced by zero.

Suppose at $t = t_i$ the optimal estimate is given by

$$x(i) = \frac{1}{2}\left\{c(i) + d(i)\right\} \tag{3.4}$$

where c(i) is vector of lower limits and d(i) is the vector of upper limits. Define

$$\Delta(i) = d(i) - \hat{x}(i). \tag{3.5}$$

Then

$$-\Delta(i) = c(i) - \hat{x}(i). \tag{3.36}$$

Finally define a set, S, consisting of m integers where the m integers are subscripts of the m measured states.

Now it is possible to proceed with the derivation. First consider determination of the absolute lower and upper bounds at $t = t_{i+1}$, c'(i+1) and d'(i+1) respectively, given the final corrected bounds and the optimal estimate at $t = t_i$. These bounds can be calculated by iterating the differences between the estimate and the final bounds at $t = t_i$ through a linear expression formed by linearizing the nonlinear system function about the nominal trajectory and subtracting and adding these quantitites to the nominal trajectory at $t = t_{i+1}$. The nominal trajectory at $t = t_{i+1}$ is found by iterating the optimal estimate at $t = t_i$ through the given nonlinear system function. Finally the effect of the input is accounted for with the resulting absolute lower limit given by

$$c'(i+1) = f\left\{\hat{x}(i)\right\} - A^{+}(i)\ \Delta(i) + A^{-}(i)\ \Delta(i) + G\underline{u} \tag{3.7}$$

and the absolute upper limit given by

$$d'(i+1) = f\{\hat{x}(i)\} + A^{+}(i)\Delta(i) - A^{-}(i)\Delta(i) + G\bar{u} \qquad (3.8)$$

Clearly the nominal portion of the trajectory in (3.7) and (3.8) is given by $f\{x(i)\}$. Also it is assumed in (3.7) and (3.8) that u has a vector of lower limits, $\underline{u}$, and a vector of upper limits, $\bar{u}$. Note that $\underline{u} = -\bar{u}$, i.e., the limits are symmetrical about zero and that the limits may vary with time as long as they are symmetrical about zero and are known at each time instant.

In (3.7) the expression $-A^{+}(i)\Delta(i) + A^{-}(i)\Delta(i)$ represents the incremental iteration below the nominal trajectory at $t = t_{i+1}$ resulting from the linearization of $f\{x(i)\}$ at $t = t_i$, A(i), and the difference between the nominal trajectory and final lower limit at $t = t_i$, $\Delta(i)$. It is possible that at $t = t_i$ certain system states may lie near their lower limit whereas others may lie near their upper limit. In order to find the absolute reachable lower limit for each state at $t = t_{i+1}$, each state must be assumed to be at the point within its bounds at $t = t_i$ which will give the maximum negative contribution to each element in c'(i+1). Since this iteration about the nominal trajectory is linear, it will be assumed that each state lies at one extremity of its bounds or the other depending upon the sign of the corresponding element in A(i). The final lower and upper bounds are symmetrical about the nominal trajectory at $t = t_i$, and the incremental iteration below the nominal trajectory used to obtain c'(i+1) is accomplished by separating A(i) into a matrix of positive values, $A^{+}(i)$, and a matrix of negative values, $A^{-}(i)$, and using $-A^{+}(i)\Delta(i) + A^{-}(i)\Delta(i)$. The incremental expression used to find d'(i+1) in (3.8)

is the same as that used in (3.7) with the sign changed.

Next it is necessary to separate the limits into two categories: those which can be affected directly by the measurement and those which cannot be affected directly by the measurement. Clearly limits of measured states can be directly affected, and limits of unmeasured states cannot be directly affected. The H matrix can be used to this end since it has "1"'s on the main diagonal for each measured state and zeros elsewhere; therefore, those limits which can be directly affected by the measurement are given by

$$\hat{c}(i+1) = Hc'(i+1) \tag{3.9}$$

and

$$\hat{d}(i+1) = Hd'(i+1). \tag{3.10}$$

Lower and upper limits which cannot be directly affected are then given by

$$c^1(i+1) = (I-H)c'(i+1) \tag{3.11}$$

and

$$d^1(i+1) = (I-H)d'(i+1) \tag{3.12}$$

where I is an n x n identity matrix.

Now it is necessary to define decision matrices which will be used to separate limits of measured variables, i.e., limits which can be directly affected by the measurement, into two categories: those

which are directly affected by the measurement and those which are not. First suppose that the corrupting noise, v(i+1), has lower and upper limit vectors given respectively by

$$-b = -\begin{vmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{vmatrix} \tag{3.13}$$

and

$$b = \begin{vmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{vmatrix} . \tag{3.14}$$

(It is possible for $b_1, b_2, \ldots, b_n$ to vary with time as long as their values are known at each time instant.) It is now possible to define decision matrices which can be used to isolate those bounds which are actually affected by the measurement.

To this end define

$$\left.\begin{aligned} k_{cj}(i+1) &= 1, \; y_j(i+1) \leq \tilde{c}_j(i+1) + b_j \\ &= 0, \text{ otherwise} \end{aligned}\right\} \; j \epsilon S$$

$$= 0, \; j \notin S \tag{3.15}$$

and

$$\left.\begin{aligned} k_{dj}(i+1) &= 1,\ y_j(i+1) \geq \tilde{d}_j(i+1) - b_j \\ &= 0,\ \text{otherwise} \end{aligned}\right\}\ j\epsilon S \qquad (3.16)$$

$$= 0,\ j\notin S.$$

These elements are used to form the main diagonals of the two decision matrices:

$$K_c(i+1) = \begin{vmatrix} k_{c1}(i+1) & & 0 \\ & \ddots & \\ 0 & & k_{cn}(i+1) \end{vmatrix} \qquad (3.17)$$

and

$$K_d(i+1) = \begin{vmatrix} k_{d1}(i+1) & & 0 \\ & \ddots & \\ 0 & & k_{dn}(i+1) \end{vmatrix} \qquad (3.18)$$

The objective in defining (3.17) and (3.18) is to place a "1" on the main diagonal of $K_c(i+1)$ for each measured state whose lower limit is not affected by $y(i+1)$ with all other elements of $K_c(i+1)$ zero and to place a "1" on the main diagonal of $K_d(i+1)$ for each measured state whose upper limit is not affected by $y(i+1)$ with all other elements of $K_d(i+1)$ zero.

At this point the various limits can be separated into three categories: those which cannot be affected by the measurement, $c^1(i+1)$ and $d^1(i+1)$, given by (3.9) and (3.10); those which can be affected by

the measurement but are not because of the particular value of the measurement, $c^2(i+1)$ and $d^2(i+1)$, given by

$$c^2(i+1) = K_c(i+1)\ c'(i+1) \tag{3.19}$$

and

$$d^2(i+1) = K_d(i+1)\ d'(i+1); \tag{3.20}$$

and those which can be affected by the measurement and are affected because of the particular value of the measurement, $\tilde{c}^3(i+1)$ and $\tilde{d}^3(i+1)$, given by

$$\tilde{c}^3(i+1) = \{I-K_c(i+1)\}H\ c'(i+1) \tag{3.21}$$

and

$$\tilde{d}^3(i+1) = \{I-K_d(i+1)\}\ Hd'(i+1). \tag{3.22}$$

Corrected values must be substituted for the limits in the third category so that the resulting third category limits are given by

$$c^3(i+1) = \{I-K_c(i+1)\}H\{y(i+1)-b\} \tag{3.23}$$

and

$$d^3(i+1) = \{I-K_d(i+1)\}H\{y(i+1)+b\}. \tag{3.24}$$

Leaving the limits in the first two categories as they were initially iterated, the three sets of limits can be combined in a single expression for the final corrected lower limit at $t = t_{i+1}$

$$c(i+1) = c^1(i+1) + c^2(i+1) + c^3(i+1) \tag{3.25}$$

and a single expression for the final corrected upper limit at $t = t_{i+1}$

$$d(i+1) = d^1(i+1) + d^2(i+1) + d^3(i+1). \tag{3.26}$$

Substituting (3.9), (3.19), and (3.23) into (3.25) and (3.10), (3.20), and (3.24) into (3.36) gives

$$c(i+1) = (I-H)\, c'(i+1) + K_c(i+1)\, c'(i+1) + \{I-K_c(i+1)\}H\{y(i+1)-b\} \tag{3.27}$$

and

$$d(i+1) = (I-H)\, d'(i+1) + K_d(i+1)\, d'(i+1) + \{I-K_d(i+1)\}H\{y(i+1)+b\}. \tag{3.28}$$

It should be noted that the decision elements defined in (3.15) and (3.16) and the corrected values for third category limits defined in (3.23) and (3.24) all follow directly from Figure 2.

The optimal estimate at $t = t_{i+1}$ is given by

$$\hat{x}(i+1) = \frac{1}{2}\{c(i+1) + d(i+1)\}. \tag{3.29}$$

Substituting (3.27) and (3.38) into (3.29) given

$$\hat{x}(i+1) = \frac{1}{2}\ (I-H)\{c'(i+1) + d'(i+1)\} + K_c(i+1)c'(i+1) + K_d(i+1)d'(i+1) + \{I-K_c(i+1)\}H\{y(i+1)-b\} + \{I-K_d(i+1)\}H\{y(i+1)+b\}. \tag{3.30}$$

Substituting (3.7) and (3.8) into (3.30) and simplifying gives

$$\begin{aligned}\hat{x}(i+1) = (I-H)f\{\hat{x}(i)\} &+ \tfrac{1}{2}\{K_d(i+1) + K_c(i+1)\}f\{\hat{x}(i)\} \\ &+ \tfrac{1}{2}\{K_d(i+1) - K_c(i+1)\}A^{+}(i)\Delta(i) \\ &- A^{-}(i)\Delta(i) + G\bar{u} \\ &+ \tfrac{1}{2}\{I-K_c(i+1)\}H\{y(i+1)-b\} \\ &+ \tfrac{1}{2}\{I-K_d(i+1)\}H\{y(i+1)+b\}.\end{aligned} \tag{3.31}$$

Equation (3.31) is the final optimal algorithm for the problem with nonlinear plant of the type defined by (3.1) and (3.2).

It should be noted that this result has several characteristics in common with the final algorithm for the estimation problem with linear plant as given by equation (2.33). The similarity in the two algorithms is explored in the following section.

## Comparison of Estimators

Several points of similarity exist between the algorithm for the linear plant (2.33) and the algorithm for the nonlinear plant (3.31). For instance, the bounds are iterated symmetrically about a nominal trajectory in the nonlinear case so that the resulting optimal estimate for unmeasured variables or measured variables whose limits are unaffected by the measurement is found simply by iterating these variables through the given nonlinear system function. Similarly in the linear case, optimal estimates for unmeasured variables and measured

variables whose limits are unaffected by the measurement are found iterating these variables through the given transition matrix. Also, in both the linear and nonlinear case the influence of the measurement on the bounds of measured variables may be characterized pictorially as shown in Figure 2.

With regard to the mathematical representations of the final algorithms for the two cases, the nonlinear case algorithm (3.31) is clearly identical to the linear case algorithm (2.33) if $\mathbf{A}(i)$ is replaced by F, $A^{+}(i)$ is replaced by $F^{+}$, and $A^{-}(i)$ is replaced by $F^{-}$. Also it should be noted that in case all of the elements of $f(\cdot)$ are linear functions of the state variables, $x(i)$, then the estimator algorithm may be found using either (3.31) or (2.33) since they yield identical algorithms for this case, i.e., the nonlinear algorithm converges to the linear algorithm when the system function is linear.

# CHAPTER IV

## SUBOPTIMAL ESTIMATOR EXAMPLES

The constraints upon which the optimal estimator is derived are so severe as to preclude widespread practical use of the estimator under optimal conditions. Therefore, it appears relevant to consider estimator operation in suboptimal environments to see if any conclusions can be reached as to particular conditions where it might be used to advantage over other popular estimators such as the Kalman estimator. In this chapter the performance of the estimator* defined in the preceeding chapters is considered relative to the Kalman estimator performance in an environment suboptimal to both estimators. In particular, relative estimator performance is observed while three major system parameters are varied: signal bandwidth, signal to corrupting noise ratio, and signal generating system input distribution.

The overall objective of the examples in this chapter is to show, both qualitatively and quantitatively, the effect of controllable parameters on estimator performance. To fulfill this objective it is necessary to determine which of the controllable parameters effect the performance of the nonlinear estimator relative to a linear estimator, in what direction the parameter must be varied to improve performance, and the extent of the improvement.

---

*This estimator is referred to as the "nonlinear estimator".

In solving practical problems the designer frequently has control over certain parameters. The parameters which can be controlled depend upon the type of application. For instance, in an open loop problem where the only objective is to design an estimator which minimizes MSE, the only parameter which can be controlled is the sample rate. On the other hand, a closed loop problem, where the objective is to design both an estimator and a controller to minimize some cost function, offers the opportunity to control parameters other than the sample rate.

An ultimate decision in choosing the nonlinear estimator over a linear estimator could be arrived at only after simulation of the particular problem of interest and comparing the nonlinear estimator with a linear estimator. Obviously there are many cases where no parameters can be controlled and the only basis for decision is simulation results. However, it is intended that the results of the examples in this chapter will provide a basis for determining the practicality of proceeding with simulations and some directions which might be pursued in such simulations when controllable parameters do exist.

The Kalman estimator is used for comparison purposes because it is well known and enjoys widespread practical use. The Kalman estimator and the nonlinear estimator are designed for mutually exclusive subsets of the set of general estimation problems, i.e., there are no estimation problems for which both are optimal. Therefore, it is necessary to choose problems which are suboptimal to both estimators.

Since the Kalman estimator is designed primarily for linear systems, or linearized versions of nonlinear systems, all the examples used are linear.

In all cases the equations used for the Kalman estimator are those found in Sage [19]. These equations are somewhat different from those found in most references in that they use the most current measurement in making the current estimate. Since these equations involve no prediction, as do equations which ignore the most current measurements, they give somewhat better results both in the steady state and transient regions.

A special note should be made in considering variations in the signal bandwidth parameter. Increasing the signal bandwidth gives the same results as decreasing the sample rate, and decreasing the signal bandwidth gives the same results as increasing the sample rate. Therefore, both of these parameter variations can be considered with the same example.

It should be remembered that in all cases, optimal or suboptimal, the nonlinear estimator is useful only if the signal generating system dynamics are bounded with known bounds.

Each of the three parameter variations are considered in separate examples in the following.

## Example One

In this example the performance of the nonlinear estimator relative to the Kalman estimator is considered as a function of the bandwidth of the signal to be estimated.

A first order, discrete system represented by

$$x(i+1) = Ax(i) + Bu(i+1) \tag{4.1}$$

is chosen as the signal generating system with the measurement given by

$$y(i+1) = x(i+1) + v(i+1). \tag{4.2}$$

The quantities u and v are chosen from uniformly distributed random ensembles such that the following expected values hold:

$$E[u(i)\ u(j)] = 0,\ i \neq j$$
$$= Q,\ i = j. \tag{4.3}$$

$$E[v(i)\ v(j)] = 0,\ i \neq j$$
$$= R,\ i = j. \tag{4.4}$$

$$E[u(i)\ v(j)] = 0, \text{ all } i \text{ and } j. \tag{4.5}$$

The system initial condition, x(0), is chosen from a uniformly distributed ensemble as are u(i) and v(i). Q is fixed at 4/3 and R is fixed at 1/3. A and B are chosen to give the desired bandwidth and also to maintain

$$E[x(i)\ x(i)] = 0, \text{ all } i. \tag{4.6}$$

The system is initialized at t = 0 and is run for an equivalent of 10 seconds real time. This means that 100 samples are taken at 10 samples/second. At each sample instant two estimates of the system state are made, one estimate using a Kalman estimator and the other estimate using the nonlinear estimator; and the actual error in each estimate is calculated. This procedure is repeated 100 times. The

MSE for both estimators is calculated at each time instant by squaring the previously calculated error figures for both estimators and then averaging across both ensembles at each time instant. This entire operation is repeated for various values of system bandwidth ranging from 100 radians/second to .0001 radians/second. The system gain is adjusted for each different system bandwidth so that the signal variance is maintained constant. In order to display the results, a fixed time instant (t = 1 second) is chosen and the MSE for both estimators at this time instant is displayed as a function of system bandwidth. (This particular time instant is rather arbitrarily chosen with only the restriction that it be in the range where both estimators have concluded their transient response and are in a steady state mode.)

The results are shown in Figure 6 and Figure 7. The low frequency end of the spectrum considered is shown in Figure 6 and the high frequency end is shown in Figure 7. For the scales used on both graphs the error without any filtering would be .33. Clearly the nonlinear estimator dominates for frequencies above approximately 6.7 radians/second and frequencies below approximately .0078 radians/second. In between these frequencies the Kalman estimator clearly excells.

The low frequency domination by the nonlinear estimator is attributable to the fact that since the system is initialized from a uniformly distributed ensemble, the signal distribution, at any fixed time, becomes more and more uniform as the system bandwidth decreases. This means that at any fixed time, the system tends toward optimum conditions for the nonlinear estimator as the bandwidth decreases.

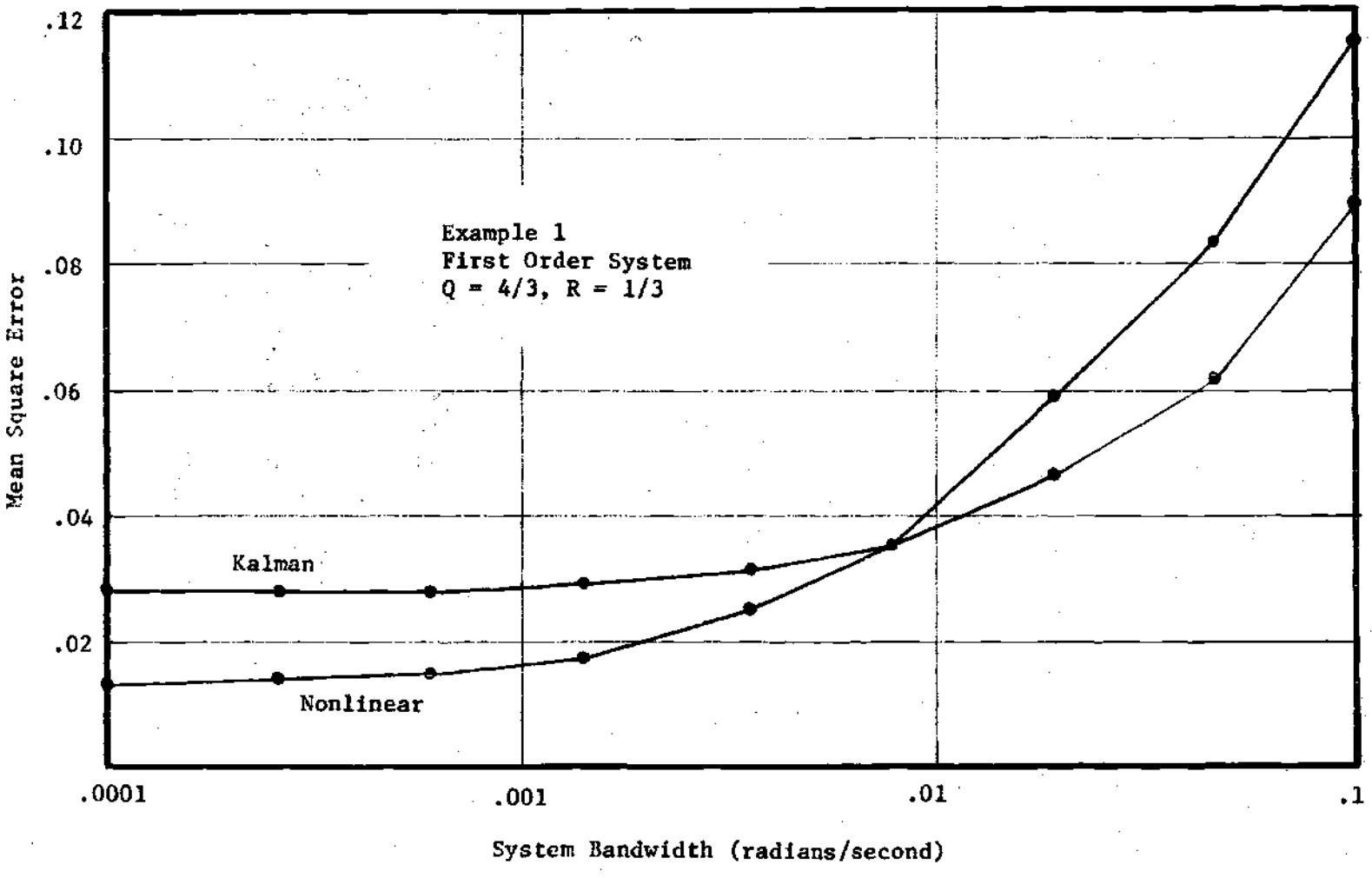



Figure 6. Steady State Estimator Response as a Function of System Bandwidth, Low Frequency Portion.

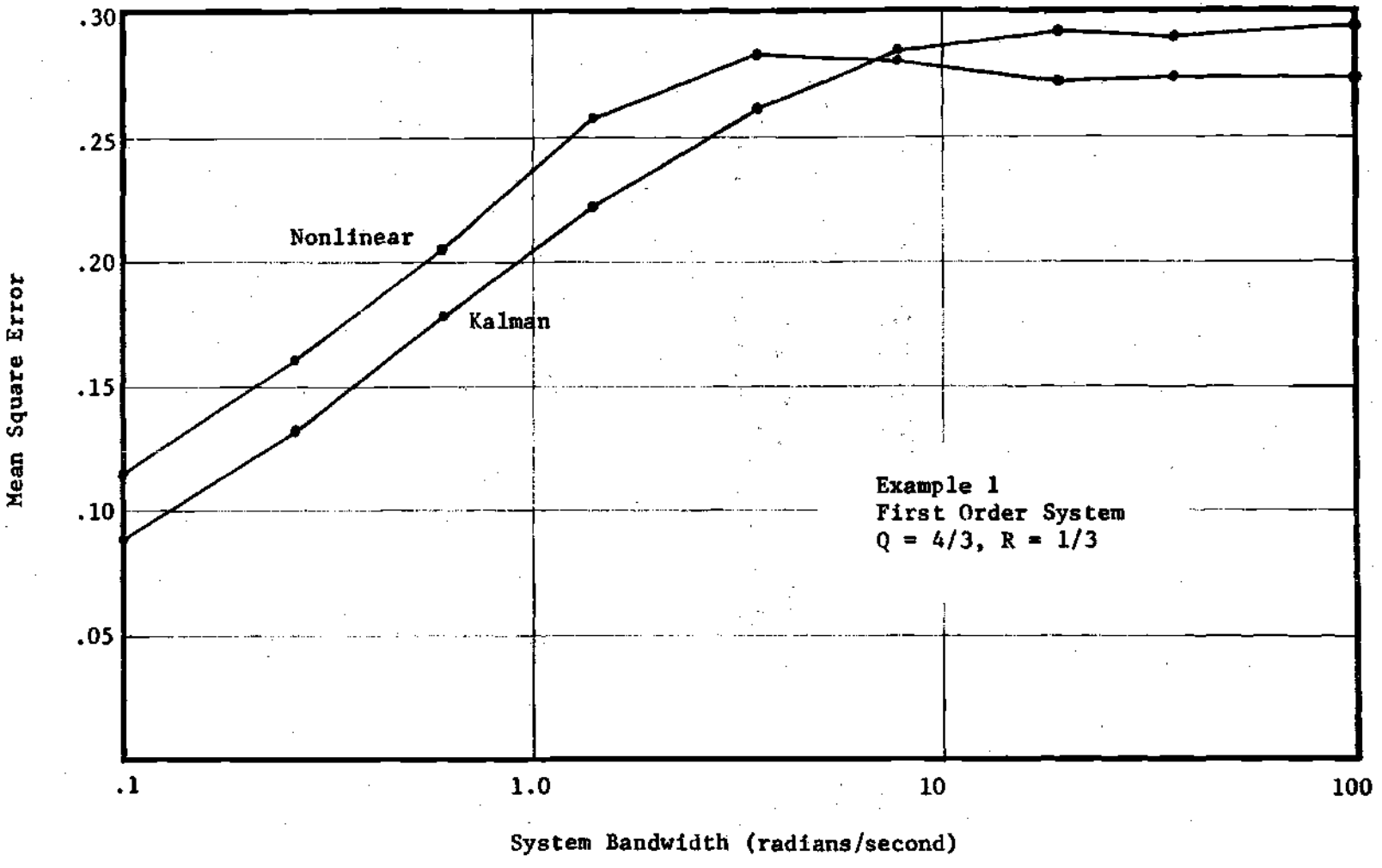



Figure 7. Steady State Estimator Response as a Function of System Bandwidth, High Frequency Portion.

The high frequency dominance of the nonlinear estimator can be explained similarly. As the system bandwidth increases the signal tends to follow the input. Therefore as the bandwidth increases the signal tends to assume the statistical characteristics of the input. The input in this case has a uniform distribution, and as the bandwidth increases the signal tends to a uniform distribution. A uniform distribution on the signal yields an optimal estimate. Consequently, as the bandwidth increases the nonlinear estimator tends toward optimal performance.

The high frequency performance is not of particular interest since this condition implies a sampling rate much lower than normally would be expected, and the estimation error in this region is essentially the same that would be achieved without any estimator at all.

The low frequency performance, however, merits additional attention since it has implications of considerable practical importance. First of all, the relative improvement in MSE gained by use of the nonlinear estimator is much greater in the low frequency end of the spectrum considered. In particular, the reduction in MSE gained by using the nonlinear estimator over the Kalman estimator is approximately 8 percent at the high frequency end, and the reduction in MSE is approximately 54 percent at the low frequency end when the nonlinear estimator is used.

Secondly, the "narrow" bandwidth of the system at the low frequency end of the spectrum is "narrow" relative to the sample rate. Therefore, in practical problems where the signal generating system bandwidth is fixed, the bandwidth can always be placed in the "narrow"

category by increasing the sampling rate. Obviously a higher sampling rate involves additional cost which must be weighed against the improvement in MSE. In cases where the signal generating system is naturally narrow, say on the order of a few radians/second, the increase in cost for a higher sampling rate is minimal, and these circumstances exist in many practical problems such as those involving vehicular dynamics.

Clearly, in cases where the reduction in error is worth the additional cost of higher sampling frequencies, the nonlinear estimator can be used to considerable advantage over the linear estimator.

To carry this example to its logical conclusion at the low frequency end, another system was constructed where the signal is fixed after being initially chosen from a uniformly distributed ensemble with known bounds and variance Q. The corrupting noise is random and at each sample instant chosen from a uniformly distributed ensemble with known bounds and variance R. The corrupting noise values are statistically independent as in all the other examples.

For this type of estimation problem both estimators will converge to the actual system state as the number of samples increases, and the relative performance must be considered in terms of rate of convergence. In order to determine if one of the estimators has distinct superiority over the other in its rate of convergence it is necessary to test both estimators with three different sets of values for signal variance, Q, and noise variance, R. This is necessary because the two estimators have different start-up procedures. For

instance, the Kalman estimator essentially ignores the first measurement and gives $\hat{x} = 0$ for its first estimate; therefore, the initial MSE, $t = 0$, is Q. On the other hand, the nonlinear estimator gives as its first estimate the actual measurement taken at $t = 0$ with the resulting MSE being R. Consequently, it is necessary to test the rate of convergence for both estimators for three cases: $Q > R$, $Q = R$, $Q < R$. In the first case the nonlinear estimator has an initial advantage; in the second case neither estimator has initial advantage; and in the third case the Kalman estimator has the initial advantage.

For each of the above three cases, the system is initiated by taking a sample at $t = 0$, and a total of 25 samples are taken. At each sample instant both estimators estimate the system state and the actual error is recorded. This procedure is repeated 100 times after which the actual MSE is computed by averaging the square of the error across the ensemble of error values at each sample instant.

The results are shown in Figures 8 through 13. The results for the first case, $Q > R$, are shown in Figure 8 and Figure 9. Q is fixed at 4/3 and R is fixed at 1/3. Figure 8 shows the MSE for both estimators at the first 10 sample points, and Figure 9 shows MSE from the ninth through the twenty-fifth sample point on an expanded scale. It is necessary to display the data on two different scales because of the wide range of error values from first to last sample point. Clearly the nonlinear estimator dominates from start-up to the last sample point.

Results for the second case, $Q = R$, are shown in Figure 10 and Figure 11. Q and R are fixed at 1/3. Once again the data is presented

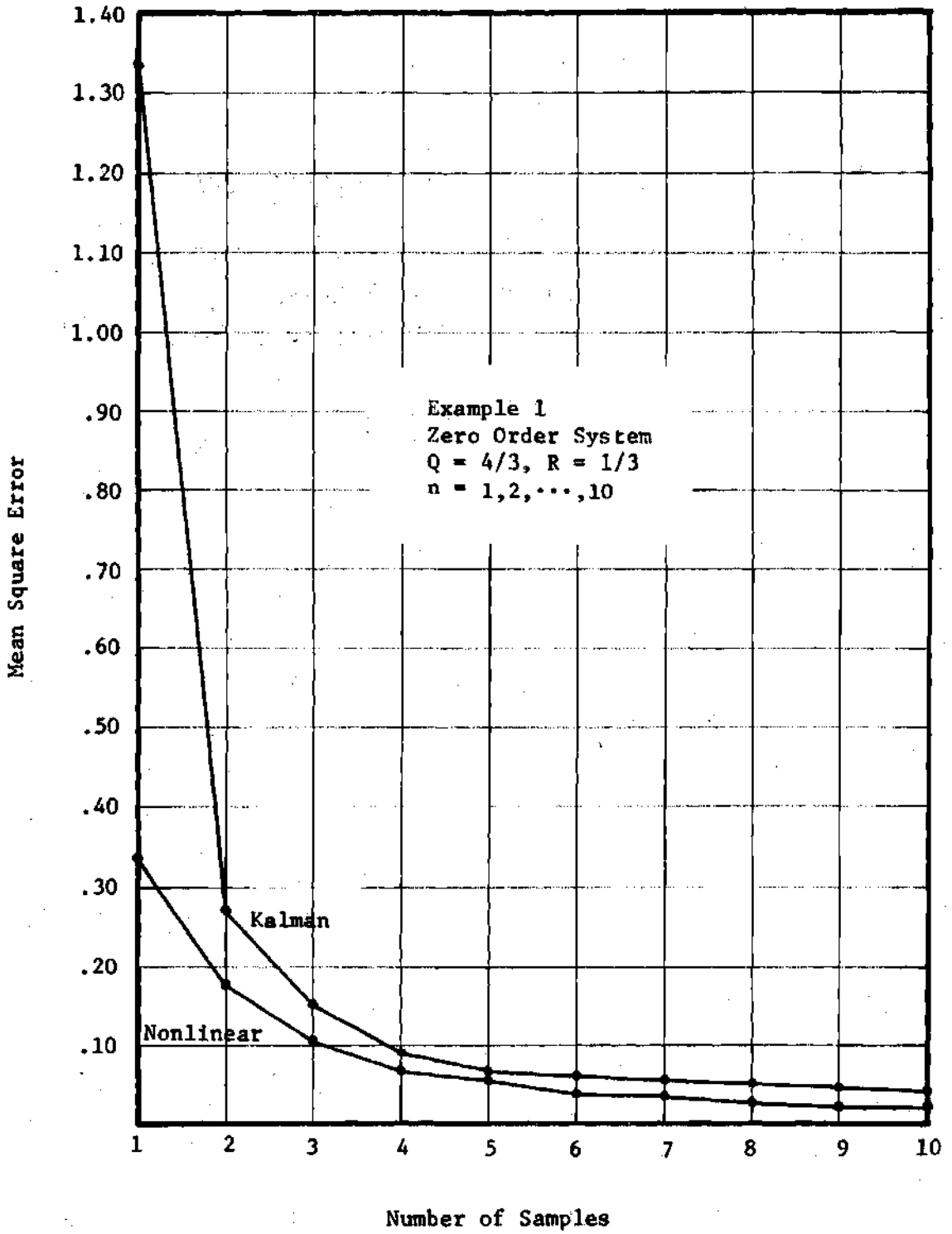



Figure 8. Estimator Response as a Function of Number of Samples, Q > R.

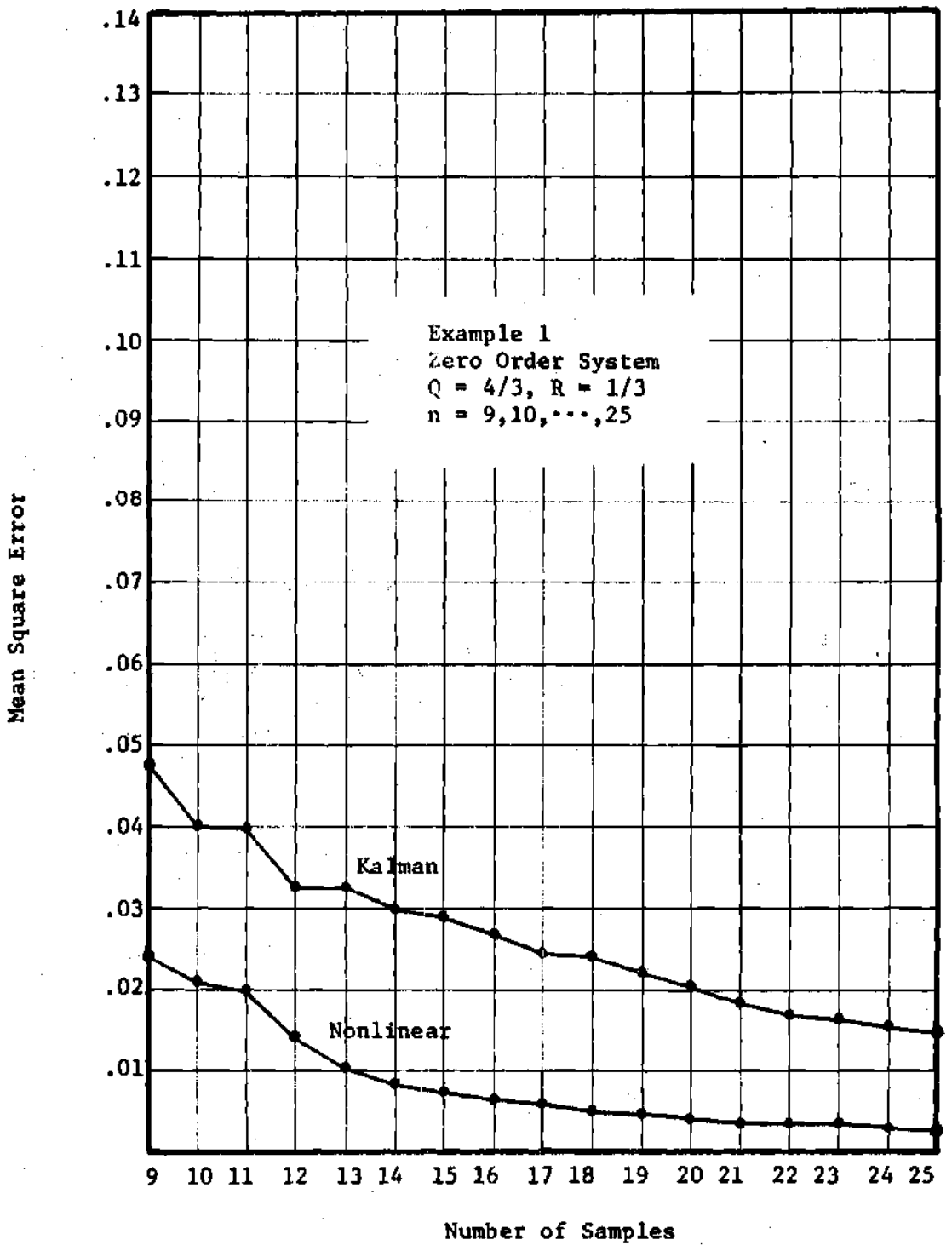



Figure 9. Estimator Response as a Function of Number of Samples, Q > R, Expanded Scale.

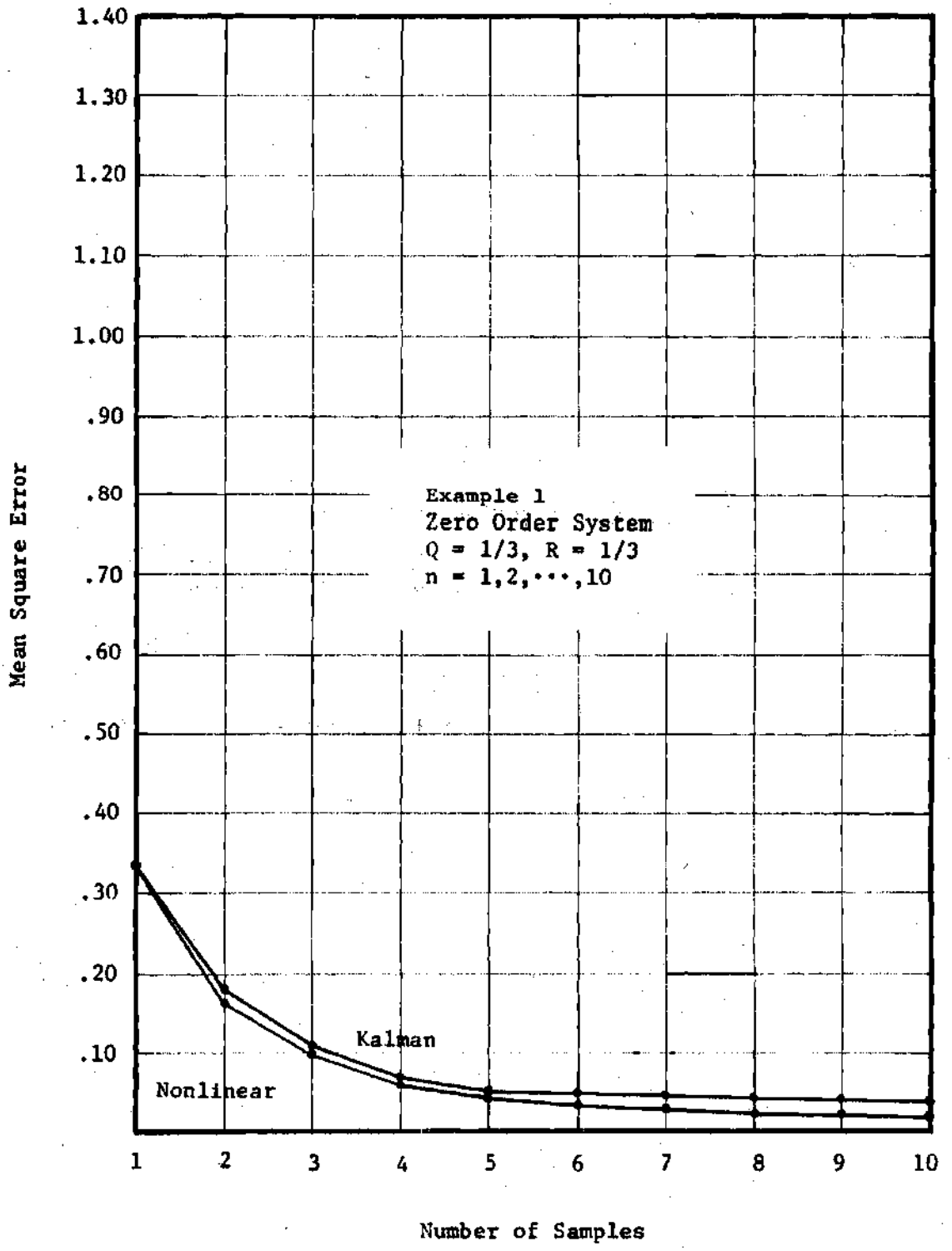



Figure 10. Estimator Response as a Function of Number of Samples, Q = R.

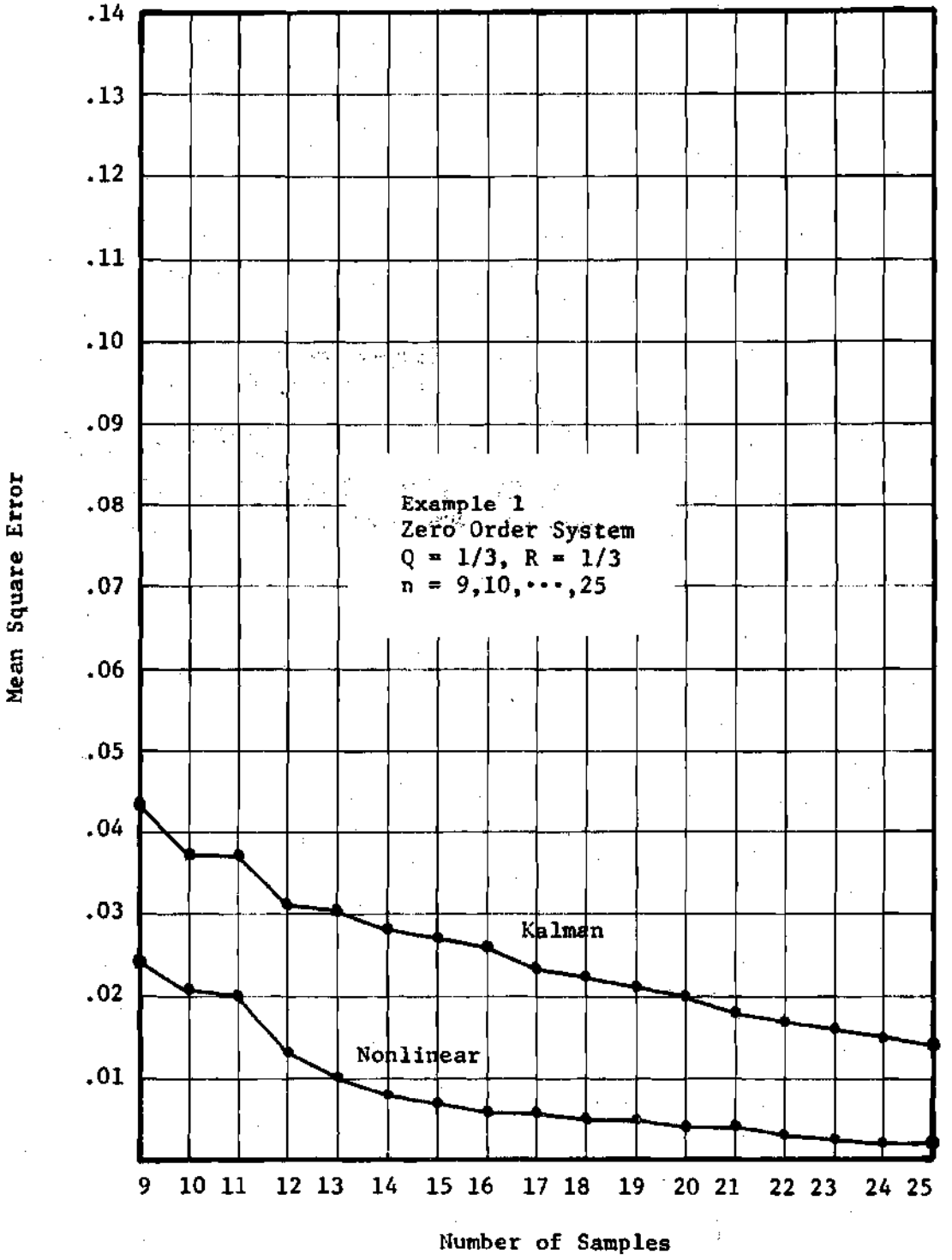



Figure 11. Estimator Response as a Function of Number of Samples, Q = R, Expanded Scale.

on two different scales because of the range of values involved. Both estimators have essentially the same MSE at the first sample instant; however, the nonlinear estimator clearly dominates from the second sample onward.

The third case results, $Q < R$, are shown in Figure 12 and Figure 13. Q is fixed at 1/3 and R is fixed at 4/3. Here again the data is presented on two different scales. The Kalman estimator clearly dominates initially because of the relative values of Q and R. This dominance of the Kalman estimator decreases and at the seventh sample point the nonlinear estimator has less MSE than the Kalman estimator. From this point on the nonlinear estimation dominates in convergence with the degree dominance continuing to increase to the last sample point.

The dominance of the nonlinear estimator in all three cases applies directly to several types of practical problems such as the estimation of fixed system parameters and the estimation of fixed system states. An example of the latter is the problem of estimating the velocity of a vehicle traveling at a fixed but unknown rate.

With regard to practical problems, two situations could exist where the nonlinear estimator might be used to advantage:

1. The cost of estimation is fixed and the objective is to achieve the greatest possible degree of accuracy.

2. The required degree of accuracy is fixed and the objective is to achieve this accuracy while minimizing the cost of estimation.

The first situation implies that the number of samples to be taken is fixed. For all three cases ($Q > R$, $Q = R$, and $Q < R$) better

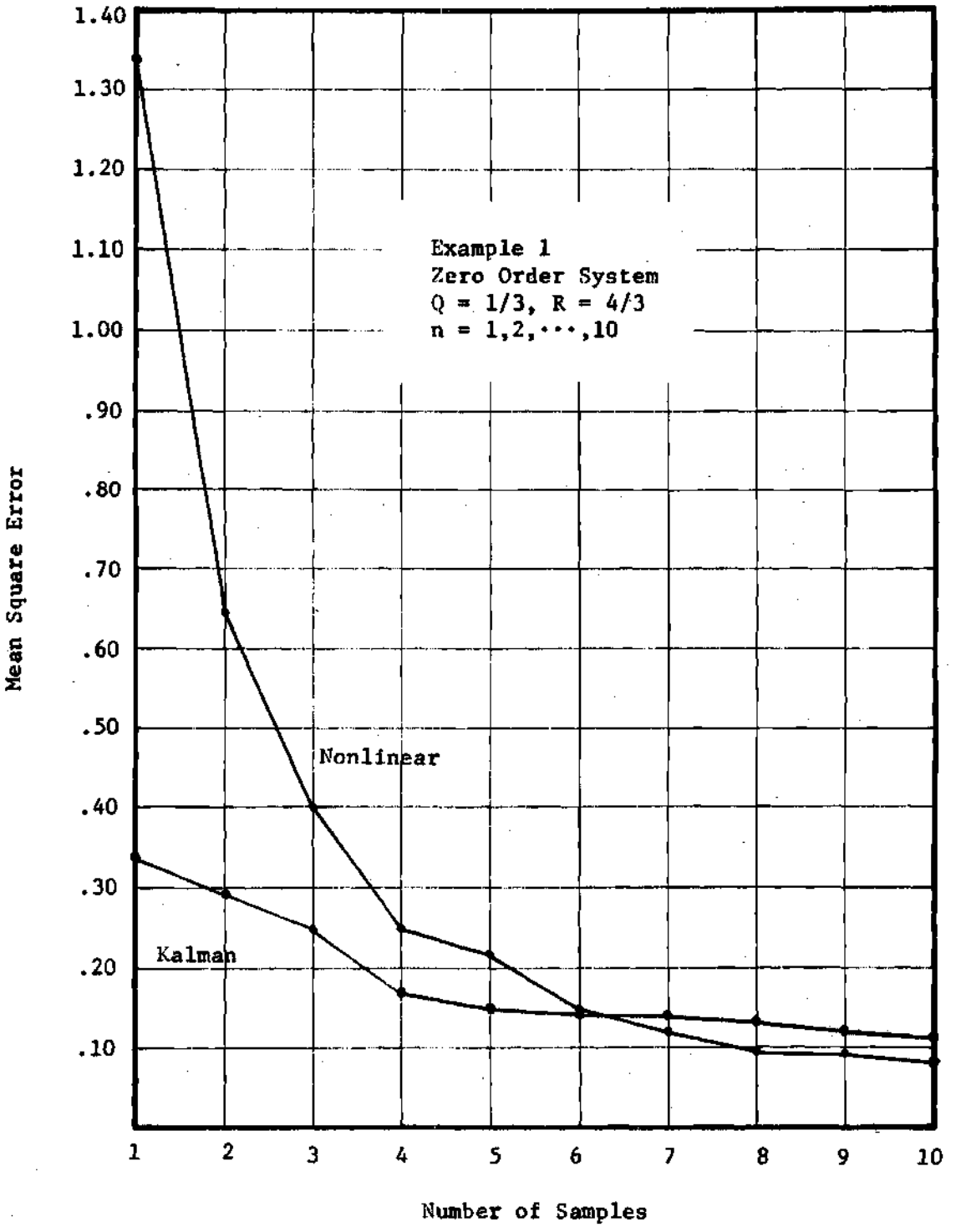



**Figure 12.** Estimator Response as a Function of Number of Samples, $Q < R$.

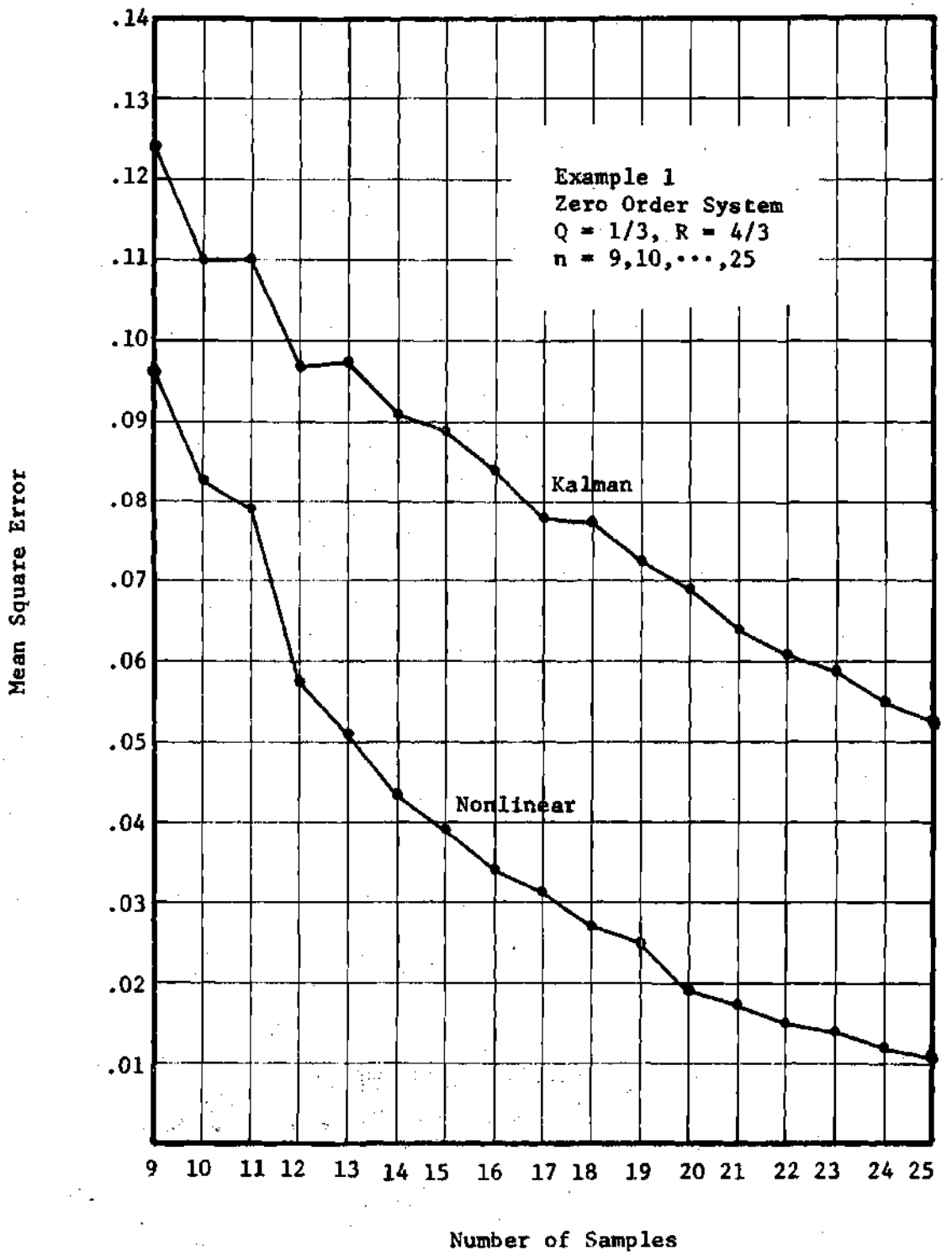



Figure 13. Estimator Response as a Function of Number of Samples, Q < R, Expanded Scale.

performance is achieved using the nonlinear estimator with the exception of the case where Q < R and the maximum number of samples which can be taken is very small. The improvement gained by using the nonlinear estimator generally increases as the number of samples increases. If 25 samples are taken, the error is reduced by approximately 83 percent for Q > R, approximately 86 percent for Q = R, and approximately 80 percent for Q < R. The percentage improvement decreases then as the number of samples decreases.

The second situation implies that the number of samples required to achieve a fixed degree of accuracy be minimized. For all three cases (Q > R, Q = R, and Q < R) the nonlinear estimator excells with the exception of the case where Q < R and the required degree of accuracy is very rough. For instance, suppose in the case where Q > R the maximum allowable MSE is .03 (roughly 10 percent of the MSE with no estimator), the minimum number of samples required using the Kalman estimator is 14 whereas the minimum number of samples required using the nonlinear estimator is 7. This implies a cost saving of 50 percent.

In order to verify the correct operation of the two Kalman estimators used in the two different parts of this example it is necessary to compare the actual MSE data found in running the example programs to the optimum MSE which would result if the Kalman estimators were operating under optimum conditions.

The first order Kalman estimator used in the first part of the example is operated using several different values of system bandwidth from 100 radians/second to .0001 radians/second. The actual MSE as a

function of time has been compared to the optimum MSE for all values of bandwidth used; however, the only comparison included is for a bandwidth of .1 radians/second, and this is shown in Table 1. The comparison data covers a total of 2.4 seconds of real time in order to display fluctuations in actual MSE in the steady state mode.

The zero order Kalman estimator used in the second part of the example is operated for three different sets of values for Q and R. For these three cases the actual MSE is compared to the optimum MSE in Table 2, Table 3, and Table 4.

All of the optimum MSE calculations have been made by hand using the Kalman estimator algorithm. Considering that the actual MSE is an average of squared error taken over a finite ensemble and that the Kalman estimators are not operating under optimum conditions, the actual MSE is remarkably close to the optimum MSE for all cases in both parts of the example.

## Example Two

In this example the performance of the nonlinear estimator relative to the Kalman estimator is considered as a function of the ratio of signal strength to corrupting noise strength.

A second order, discrete system represented by

$$x(i+1) = Ax(k) + Bu(i+1) \tag{4.7}$$

is chosen as a signal generating system. The measurement is given by

$$y(i+1) = Hx(i+1) + v(i+1) \tag{4.8}$$

Table 1. Actual and Optimum MSE for First Order Kalman Estimator, Example 1.

| Time (Seconds) | Optimum MSE | Actual MSE |
|---|---|---|
| 0.0 | 1.3333 | 1.4605 |
| 0.1 | .2667 | .2709 |
| 0.2 | .1544 | .1656 |
| 0.3 | .1159 | .1255 |
| 0.4 | .0986 | .0946 |
| 0.5 | .0899 | .0727 |
| 0.6 | .0852 | .0683 |
| 0.7 | .0827 | .0708 |
| 0.8 | .0813 | .0845 |
| 0.9 | .0804 | .0856 |
| 1.0 | .0800 | .0883 |
| 1.1 | .0797 | .0847 |
| 1.2 | .0796 | .0850 |
| 1.3 | .0795 | .0782 |
| 1.4 | .0795 | .0924 |
| 1.5 | .0794 | .0998 |
| 1.6 | .0794 | .0961 |
| 1.7 | .0794 | .0826 |
| 1.8 | .0794 | .0954 |
| 1.9 | .0794 | .0794 |
| 2.0 | .0794 | .0834 |
| 2.1 | .0794 | .0782 |
| 2.2 | .0794 | .0700 |
| 2.3 | .0794 | .0683 |
| 2.4 | .0794 | .0700 |

Table 2. Actual and Optimum MSE for Zero Order Kalman Estimator, Example 1, Q > R.

| Number of Samples | Optimum MSE | Actual MSE |
|---|---|---|
| 1 | 1.3333 | 1.3633 |
| 2 | .2667 | .2743 |
| 3 | .1481 | .1524 |
| 4 | .1026 | .0904 |
| 5 | .0784 | .0682 |
| 6 | .0635 | .0620 |
| 7 | .0533 | .0553 |
| 8 | .0460 | .0503 |
| 9 | .0404 | .0475 |
| 10 | .0360 | .0401 |
| 11 | .0325 | .0397 |
| 12 | .0296 | .0329 |
| 13 | .0272 | .0330 |
| 14 | .0252 | .0297 |
| 15 | .0234 | .0288 |
| 16 | .0219 | .0269 |
| 17 | .0205 | .0246 |
| 18 | .0193 | .0239 |
| 19 | .0183 | .0219 |
| 20 | .0173 | .0205 |
| 21 | .0165 | .0185 |
| 22 | .0157 | .0171 |
| 23 | .0150 | .0165 |
| 24 | .0143 | .0154 |
| 25 | .0137 | .0147 |

Table 3. Actual and Optimum MSE for Zero Order Kalman Estimator, Example 1, Q = R.

| Number of Samples | Optimum MSE | Actual MSE |
|---|---|---|
| 1 | .3333 | .3408 |
| 2 | .1667 | .1804 |
| 3 | .1111 | .1171 |
| 4 | .0833 | .0719 |
| 5 | .0667 | .0576 |
| 6 | .0556 | .0556 |
| 7 | .0476 | .0492 |
| 8 | .0417 | .0451 |
| 9 | .0370 | .0429 |
| 10 | .0333 | .0367 |
| 11 | .0303 | .0365 |
| 12 | .0278 | .0308 |
| 13 | .0256 | .0309 |
| 14 | .0238 | .0281 |
| 15 | .0222 | .0273 |
| 16 | .0208 | .0255 |
| 17 | .0196 | .0234 |
| 18 | .0185 | .0229 |
| 19 | .0175 | .0211 |
| 20 | .0167 | .0198 |
| 21 | .0159 | .0180 |
| 22 | .0152 | .0169 |
| 23 | .0145 | .0162 |
| 24 | .0139 | .0152 |
| 25 | .0133 | .0144 |

Table 4. Actual and Optimum MSE for Zero Order Kalman Estimator, Example 1, Q < R.

| Number of Samples | Optimum MSE | Actual MSE |
|---|---|---|
| 1 | .3333 | .3408 |
| 2 | .2667 | .2910 |
| 3 | .2222 | .2367 |
| 4 | .1905 | .1674 |
| 5 | .1667 | .1482 |
| 6 | .1481 | .1423 |
| 7 | .1333 | .1372 |
| 8 | .1212 | .1285 |
| 9 | .1111 | .1241 |
| 10 | .1026 | .1103 |
| 11 | .0952 | .1107 |
| 12 | .0889 | .0972 |
| 13 | .0833 | .0979 |
| 14 | .0784 | .0913 |
| 15 | .0741 | .0893 |
| 16 | .0702 | .0838 |
| 17 | .0667 | .0783 |
| 18 | .0635 | .0779 |
| 19 | .0606 | .0725 |
| 20 | .0580 | .0693 |
| 21 | .0556 | .0641 |
| 22 | .0533 | .0614 |
| 23 | .0513 | .0589 |
| 24 | .0494 | .0553 |
| 25 | .0476 | .0525 |

In particular the signal generating system and measurement equation have the following structure:

$$x(i) = \begin{vmatrix} x_1(i) \\ x_2(i) \end{vmatrix} \tag{4.9}$$

$$A = \begin{vmatrix} .99884 & .09512 \\ -.02283 & .90371 \end{vmatrix} \tag{4.10}$$

$$B = \begin{vmatrix} .00484 \\ .09512 \end{vmatrix} \tag{4.11}$$

$$H = \begin{vmatrix} 1 & 0 \end{vmatrix} \tag{4.12}$$

with u, y, and v being scalar. Samples are taken at the rate of 10/second and u and v are generated as in Example One.

The system and the estimators are operated for 10 seconds (100 samples) with both estimators making an estimate at each sample point and the actual error for both estimates recorded. This operation is repeated 100 times with the MSE calculated as an average across the 100 point ensemble.

In the first part of this example Q and R are initially fixed at 1/3 and the MSE is recorded at a representative point in the steady state region. This procedure is repeated several times as Q is increased from 1/3 to 81/3. The results are shown pictorially in Figure 14 and Figure 15. Figure 14 shows position error for both estimators

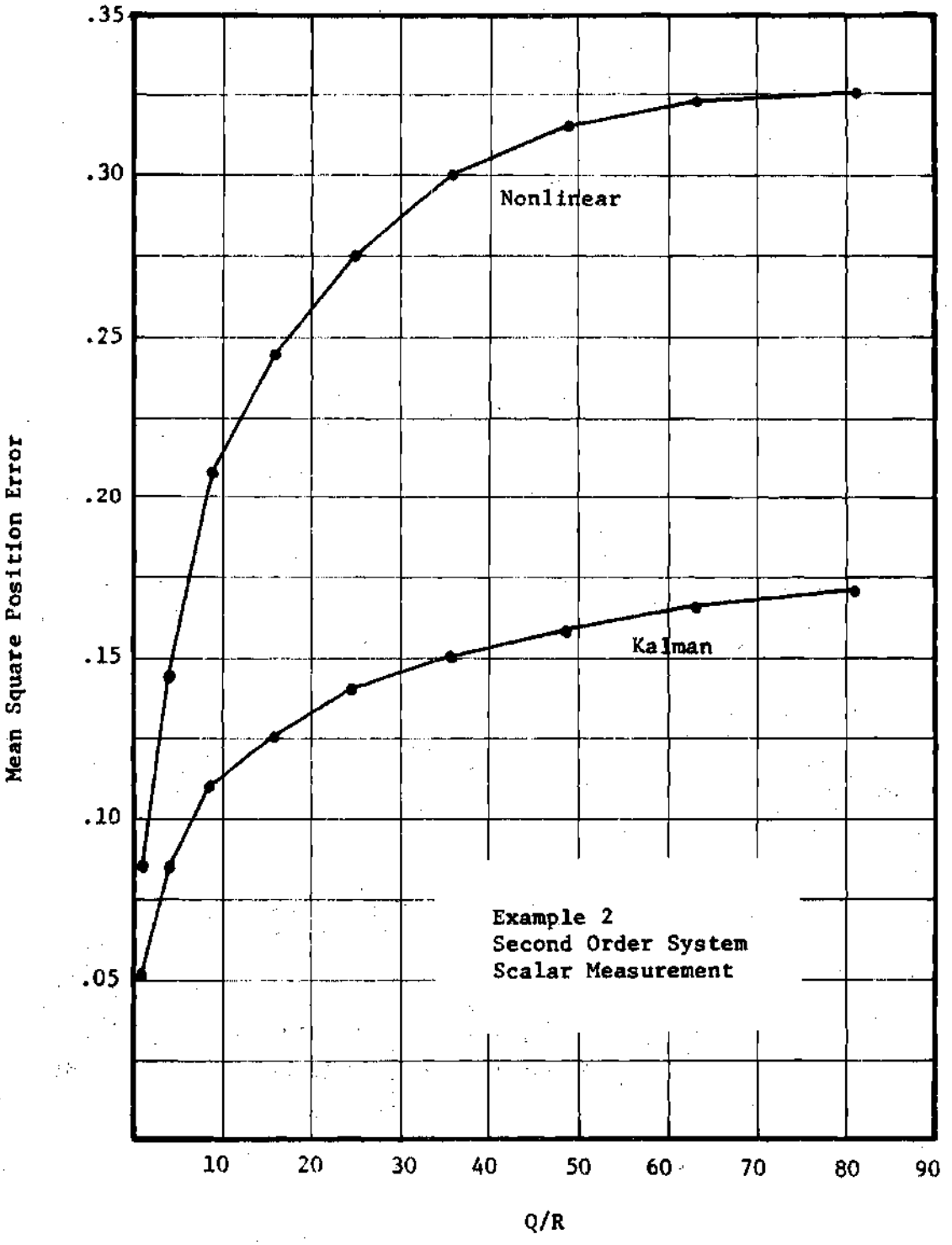



Figure 14. Steady State Position Error as a Function of Q/R, R Fixed.

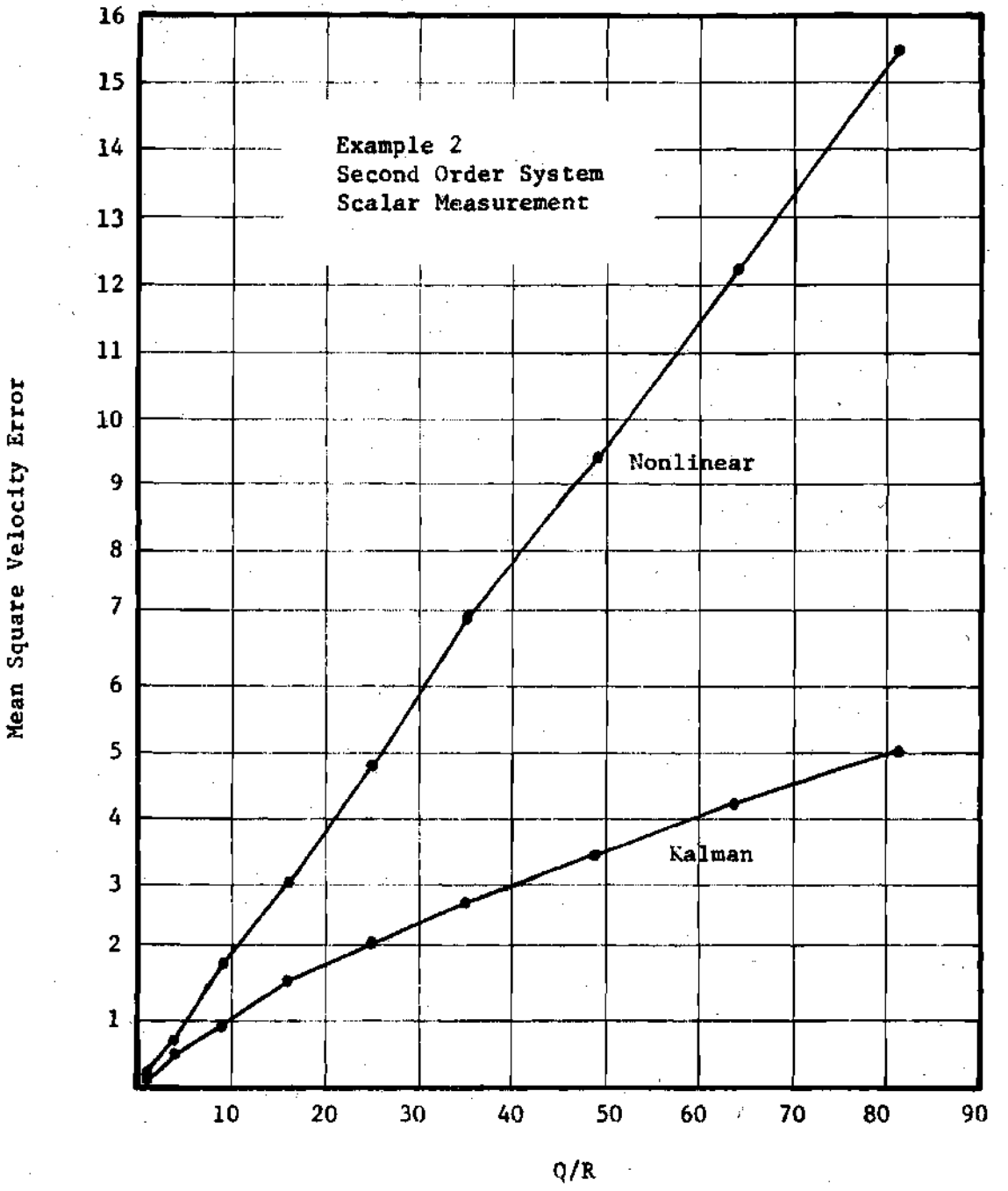



Figure 15. Steady State Velocity Error as a Function of Q/R, R Fixed.

as a function of Q/R, and Figure 15 shows velocity error for both estimators as a function of Q/R.

In the second part of this example Q and R are initially fixed at 1/3 and MSE data is taken as R is increased from 1/3 to 81/3. The results are shown in Figure 16 and Figure 17. Figure 16 shows position error as a function of R/Q, and Figure 17 shows velocity error as a function of R/Q.

Clearly the relative performance of the nonlinear estimator and the Kalman estimator is not affected by the relative strength of Q and R. This appears to be a logical conclusion since varying Q and R has no effect on the signal probability distribution; however, this result means that signal strength relative to corrupting noise strength cannot be used as a criterion to decide whether to use the nonlinear estimator or a linear estimator. Nevertheless, several interesting observations can be made on the results of this example. For instance, in the case where R is fixed and Q is increased, the system eventually reaches a point where there is very little error in the measurement as compared to the signal variance and both estimators offer very little reduction in error. However, the Kalman estimator does offer more improvement than the nonlinear estimator. It is assumed from the data that as Q increases both estimators will ultimately show negligible improvement in error. On the other hand, velocity error appears to increase in both estimators without asymptotic limit as Q increases. This occurs because the velocity state is not measured and the velocity error increases with increasing velocity variance which is linearly related to Q.

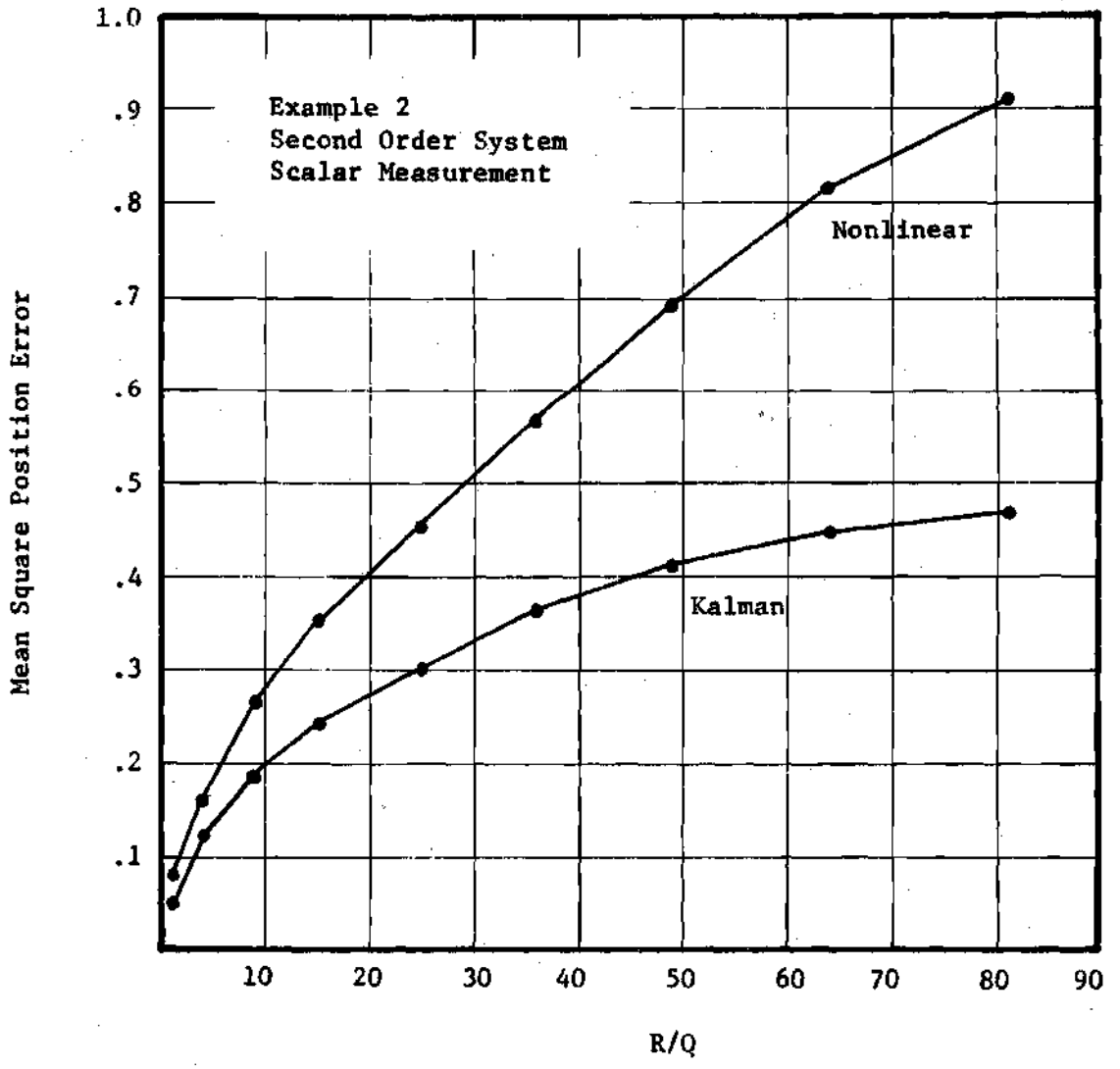



Figure 16. Steady State Position Error as a Function of R/Q, Q Fixed.

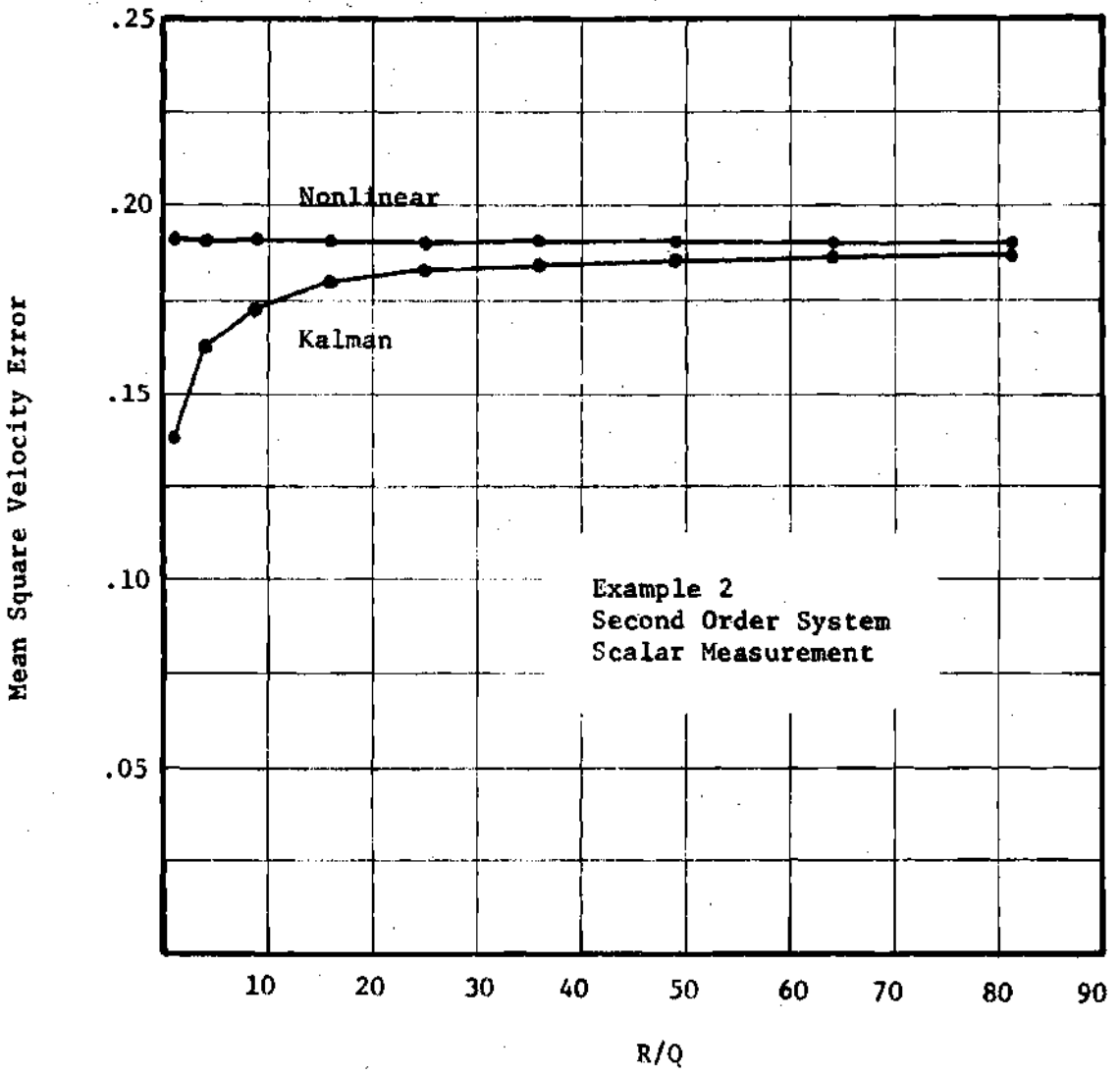



Figure 17. Steady State Velocity Error as a Function of R/Q, Q Fixed.

For the case where Q is fixed and R is increased, the position error continues to increase with R for both estimators. The velocity error for the Kalman estimator appears to asymptotically approach the fixed velocity error of the nonlinear estimator as R increases.

Once again the correct operation of the Kalman estimator is verified by calculating optimum MSE and comparing it with the actual MSE encountered in running the example program. In this example the optimum MSE calculations are performed with the aid of a computer. Comparison have been made for all values of Q and R used in the example; however, the only comparison included, Table 5, is for the case where $Q = R = 1/3$. In all cases considered the actual MSE is remarkably close to the optimum MSE. Note that the comparison data only covers 2.4 seconds of real time since the estimator is at steady state to three significant digits after 2.4 seconds.

## Example Three

In this example the performance of the nonlinear estimator relative to the Kalman estimator is observed as a function of the probability distribution of the signal forcing the signal generating system.

The signal generating and measurement systems used here are identical to those second order systems used in the previous example except that in this case Q is fixed at 1/3 and R is fixed at 1/3. The probability distribution of u is varied starting with a uniform distribution and gradually moving the weight of the distribution toward the boundaries until the distribution is represented by an impulse at either boundary, i.e., a zero mean random bang-bang input.

Table 5. Actual and Optimum MSE for Second Order Kalman Estimator, Example 2, Q = R.

| Time (Seconds) | Optimum Position Error | Actual Position Error | Optimum Velocity Error | Actual Velocity Error |
|---|---|---|---|---|
| 0.0 | .6950 | .6950 | .1667 | .1667 |
| 0.1 | .2253 | .2398 | .1666 | .1543 |
| 0.2 | .1348 | .1437 | .1662 | .1495 |
| 0.3 | .0972 | .1120 | .1649 | .1406 |
| 0.4 | .0776 | .0956 | .1629 | .1222 |
| 0.5 | .0662 | .0900 | .1601 | .1306 |
| 0.6 | .0594 | .0876 | .1567 | .1201 |
| 0.7 | .0554 | .0776 | .1528 | .1318 |
| 0.8 | .0531 | .0672 | .1487 | .1265 |
| 0.9 | .0519 | .0626 | .1447 | .1530 |
| 1.0 | .0512 | .0668 | .1409 | .1479 |
| 1.1 | .0510 | .0669 | .1375 | .1261 |
| 1.2 | .0509 | .0639 | .1345 | .1368 |
| 1.3 | .0509 | .0608 | .1320 | .1287 |
| 1.4 | .0509 | .0591 | .1300 | .1614 |
| 1.5 | .0509 | .0562 | .1284 | .1466 |
| 1.6 | .0508 | .0443 | .1271 | .1212 |
| 1.7 | .0507 | .0408 | .1262 | .1194 |
| 1.8 | .0506 | .0445 | .1256 | .1181 |
| 1.9 | .0505 | .0414 | .1251 | .1104 |
| 2.0 | .0504 | .0374 | .1248 | .0992 |
| 2.1 | .0503 | .0394 | .1246 | .1140 |
| 2.2 | .0502 | .0362 | .1245 | .1360 |
| 2.3 | .0501 | .0395 | .1245 | .1082 |
| 2.4 | .0500 | .0373 | .1244 | .1193 |

The results of this example are shown in Figure 18 and Figure 19. The data is obtained by operating the system for 2.5 seconds and taking a total of 26 samples, at 10 samples/second, counting the sample taken at t = 0. Estimates are made and the actual error is recorded at each sample point. This operation is repeated 100 times and the data shown is the result of averaging the squared error at each sample point across the entire ensemble of 100 error calculations.

The nonlinear estimator performance continues to improve as the weight of the input distribution function is moved toward the extremities. Since the total improvement in going from a uniform distribution to an impulse distribution is in the neighborhood of 40 percent and since the error curves contain a certain amount of randomness due to the use of a finite ensemble, only two error curves are shown for the nonlinear estimator. These two curves, however, show the results for the two extreme cases used in the input distribution function. The Kalman estimator shows very little change in MSE as the input distribution function is varied. Therefore, the single Kalman error curve shown for comparison with the nonlinear estimator is for the impulse input distribution case.

The improvement in the nonlinear estimator performance as the weight of the input signal distribution moves toward the boundaries is attributed to the fact that moving distribution weight to the boundaries on u(i) makes the signal distribution more uniform at any fixed point in time and, therefore, closer to optimal conditions for the nonlinear estimator. However, as time increases this advantage disappears

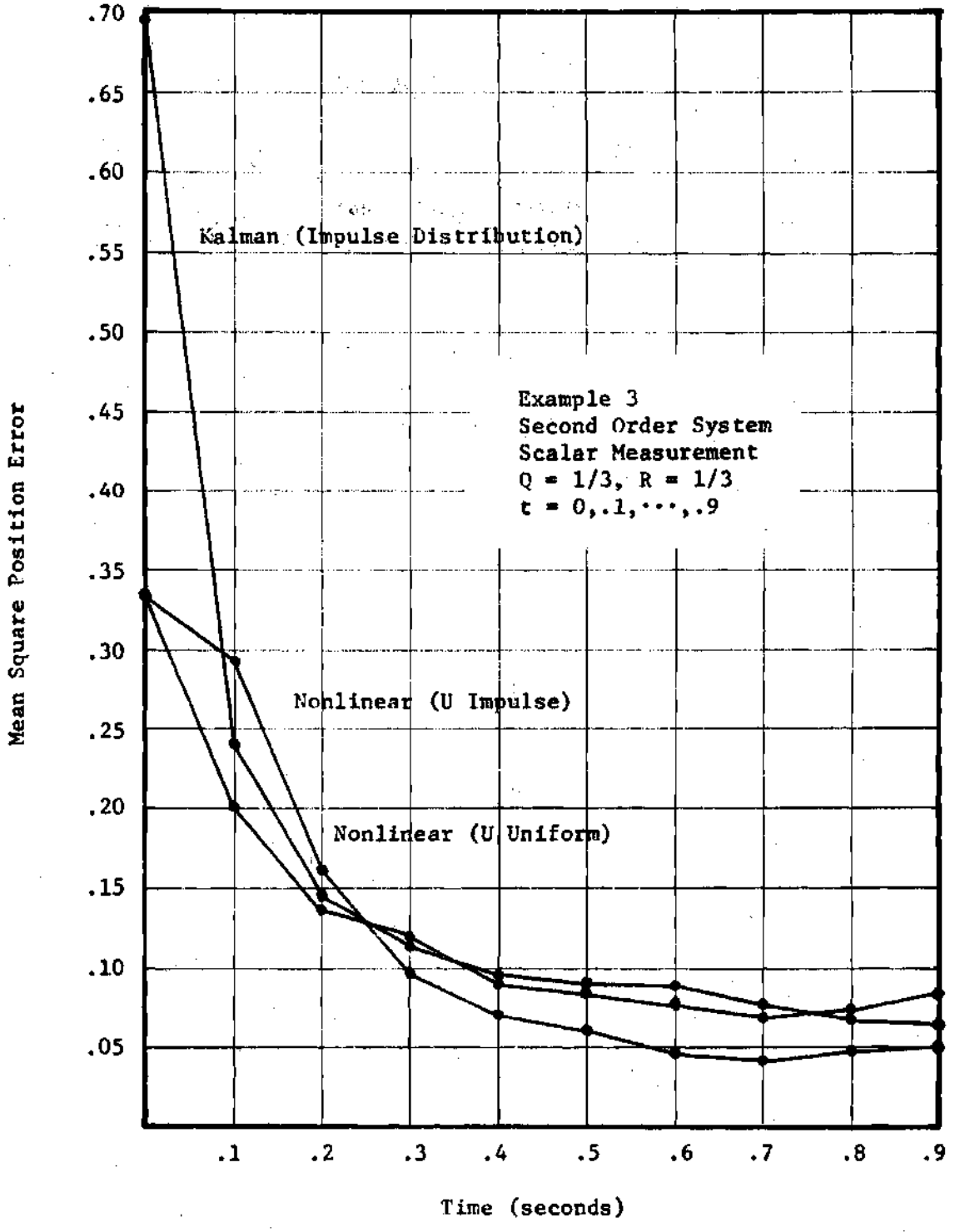



Figure 18. Estimator Transient Response for Various Input Distribution Functions.

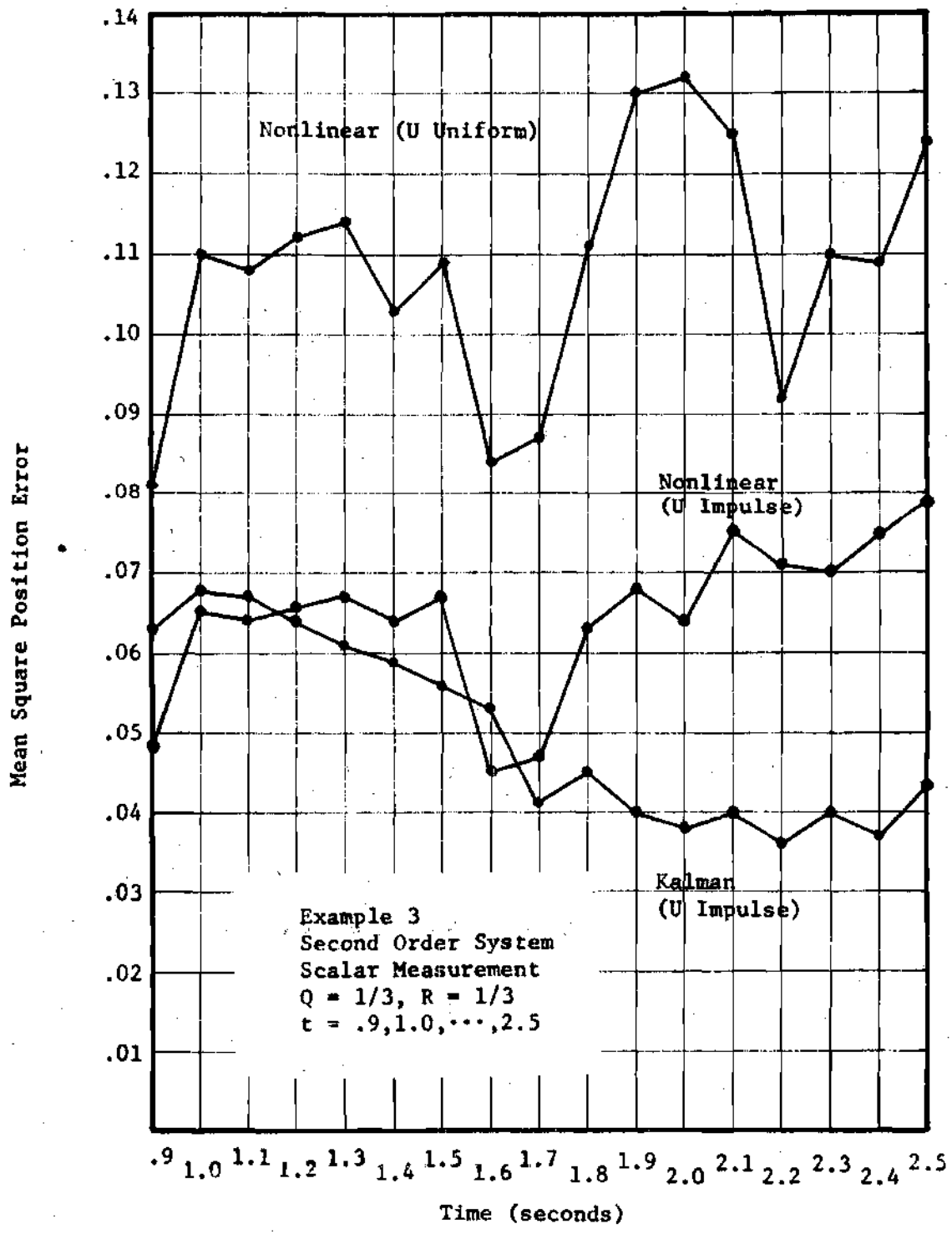



Figure 19. Estimator Transient Response for Various Input Distribution Functions, Expanded Scale.

as the input distribution has less effect on signal distribution.

For a short period of time after initiation of the estimator the nonlinear estimator performance exceeds that of the Kalman estimator for distributions on u(i) close to the impulse case. This leads to the possibility of a practical application. Consider a closed loop control-estimation problem where it is necessary for control purposes to use a bang-bang controller. Use of such a controller could lead to a case where better estimation is obtained using a nonlinear estimator rather than a linear estimator and particularly if the system is designed to operate primarily in the transient region such as the case of a reset controller used in guidance systems.

The second order Kalman estimator used in this example is identical to that used in the previous example and the correct operation of this estimator is verified by the data shown in that example in Table 5.

# CHAPTER V

## CONCLUSIONS AND RECOMMENDATIONS

In conclusion there are two major contributions inherent in this work. The first is the concept used to attack the derivation of the optimal estimator. The importance of this concept is evidenced by the fact that preliminary doctoral work in this area, transmitted by private communication, has already resulted in considerable use of this concept such as Kuo [17], Kuo and Rowland [22], and Clark [24]. The second contribution is the determination of specific areas of suboptimal application where the nonlinear estimator can be used to advantage over existing linear estimators.

The uniqueness of the derivation approach is the marriage of two concepts previously used separately: 1) the concept of the optimal estimate lying in the intersection of the set of reachable states and the set of states which could have given the observed measurement, and 2) the well-known Bayesian concept of the optimal estimate being the mean of the signal density conditioned on the measurement. Neither of these two concepts lead directly to an estimator algorithm given a specific estimation problem even if the signal generating system and signal and corrupting noise are all well defined. However, this work shows that when the two concepts are combined and certain restrictions are placed on the signal and corrupting noise, an optimal estimation algorithm can be determined in a straightforward manner for any specific

problem which meets the given restrictions.

The second contribution, determination of areas of suboptimal application, is a necessary extension of the first contribution since there are few practical cases where actual optimal conditions are exactly met; yet, there are many cases where conditions are close enough to optimal to allow suboptimal performance which will exceed that of linear estimators. In any case, optimal or suboptimal, the nonlinear estimator is useful only if the signal generating system dynamics are known and the signal generating system input and corrupting noise are bounded with known bounds. Given this constraint the examples in Chapter IV indicate at least two sets of conditions where the nonlinear estimator can be used to advantage over the Kalman estimator.

The first set of conditions exist when the additional cost of a higher sampling rate is justified by the increase in estimator performance. Example One of Chapter IV shows that as the sampling rate increases the estimator error for both the nonlinear estimator and Kalman estimator decreases; however, the nonlinear estimator error decreases much more rapidly than the Kalman estimator error. The second part of Example One carries this condition to its extremity by comparing estimator performances when the signal is fixed but unknown. For all cases considered under this condition, the nonlinear estimator converges to the true system state much faster than the Kalman estimator. In all fairness it should be noted that there exists a third alternative to this particular estimation problem: namely, estimating the system state by taking a simple average of the

measurement. For this estimation procedure the MSE is R/n, n = 1, 2, ... where R is the corrupting noise variance and n is the number of measurements used. When n = 1, the averaging estimator error is the same as the nonlinear estimator; however, the nonlinear estimator converges to the actual system state much faster than the averaging estimator as subsequent samples are taken. Further, it should also be noted that this fixed state condition exists approximately when a system is slowly varying and several rapid measurements are made at each "point" where the system state is to be estimated.

The second set of conditions exist when the signal generating system input is bang-bang and the estimator transient performance is of primary concern. Example Three shows that when the control input is bang-bang the nonlinear estimator performance exceeds that of the linear estimator for a short period of time after initiation of the estimator. Such a situation could be further enhanced in favor of the nonlinear estimator if the signal variance is much larger than the corrupting noise variance since, as previously mentioned, the initial error in the linear estimator is equal to the signal variance, and the initial error in the nonlinear estimator is equal to the noise variance. In fact, the transient performance of the nonlinear estimator will exceed that of the linear estimator when the signal variance is much larger than the corrupting noise variance regardless of the control input density. Many closed loop control-estimation design problems involve estimation where the entire control-estimation portion of the system is periodically reset such as in certain navigation problems.

In these cases the designer has some freedom in choosing the control input density, and assuming there are no physical or theoretical restrictions to prevent use of bang-bang control, he may choose this type of control and thereby create a situation where better estimation is obtained using the nonlinear estimator.

Two other conclusions should be noted:

1. As shown in Chapter II it is possible to calculate an "on-line" estimate of MSE at each sample point. Unlike the Kalman estimator it is not possible to precalculate MSE. However, the "on-line" MSE calculation gives a good idea of the accuracy of each estimate whereas the precalculated MSE of the Kalman estimator is a true average, and the error in individual estimates may fall well outside the average. The 'on-line' MSE estimate varies over a wide range on a given run and does not give a smooth curve since it is directly related to the actual accuracy of the estimate. A smooth curve may be obtained by making several runs and averaging MSE calculations across the ensemble at each sample point.

2. The nonlinear estimator is in some sense adaptive in that the major parameters, the decision matrices, are recalculated at each sample point according to the actual measurement taken. The major parameters of linear estimators are all precalculated. This precalculation can lead to errors if the actual signal generating system deviates somewhat from its assumed structure.

The work described in this dissertation might be extended in several different directions, all having the same objective of providing

a wider range of applicability. These several different directions might be roughly divided into two categories: refinement of the basic algorithm and exploration of practical applications.

Several particular problems might be attacked in refining the basic algorithm. The two most important of these are:

1. The design of additional feedback loops in the estimator to automatically correct for errors in the initial assumptions and consequently make the estimator less sensitive to such errors.

2. Translation of the restrictions which have been placed on the signal into a set of restrictions on the control input and signal generating system with the end result being the definition of a class of signal generating systems and a class of control inputs which would allow the conditional mean to be calculated using only the endpoints of the distributions involved.

# APPENDIX

## NATURE OF SIGNAL DENSITY CONDITIONED ON THE MEASUREMENT

First consider a state, say $x_j(i+1)$, which is unmeasured. Estimates for unmeasured variables at $t_{i+1}$ are obtained by iterating $\hat{x}(i)$ through the system dynamics; therefore, the statistical nature of these variables is of no concern.

Next consider a state, say $x_k(i+1)$, which is measured with measurement $y_i(i+1)$ where

$$y_k(i+1) = x_k(i+1) + v_k(i+1) \tag{A.1}$$

and $v_k(i+1)$ is the corrupting noise. (For convenience of notation in the following, the quantities $x_k(i+1)$, $y_k(i+1)$, and $v_k(i+1)$ will be referred to respectively as x, y, and v.) Considering the reachable states of x at $t_{i+1}$ it is assumed that x has a lower bound c and an upper bound d. The corrupting noise v is also bounded with a lower bound -b and an upper bound b.

Since both x and v are uniformly distributed

$$\begin{aligned} p_x(x) &= \frac{1}{d-c}, \quad c \leq x \leq d \\ &= 0, \text{ otherwise} \end{aligned} \tag{A.2}$$

$$\begin{aligned} p_v(v) &= \frac{1}{2b}, \quad -b \leq v \leq b \\ &= 0, \text{ otherwise} \end{aligned} \tag{A.3}$$

The probability density for y is found using (A.1) and the fact that x and v are statistically independent for all t:

$$p_y(y) = \int_{-\infty}^{\infty} p_x(x)\ p_v(y-x)dx. \tag{A.4}$$

Substituting (A.2) and A.3) into (A.4) and simplifying

$$p_y(y) = \frac{1}{2b}\ [p_x(y+b) - p_x(y-b)]. \tag{A.5}$$

In order to use Bayes' Rule to find $p_{x|y}(x|y)$ it is first necessary to find $p_{y|x}(y|x)$ which clearly is given by

$$\begin{aligned} p_{y|x}(y|x) &= \frac{1}{2b},\ x - b \leq y \leq x + b \\ &= 0, \text{ otherwise.} \end{aligned} \tag{A.6}$$

Now it follows directly from Bayes' Rule that

$$p_{x|y}(x|y) = \frac{p_{y|x}(y|x)\ p_x(x)}{p_y(y)} \tag{A.7}$$

Substituting (A.2), (A.5), and (A.6) into (A.7) yields

$$\begin{aligned} p_{x|y}(x|y) &= \frac{1}{(d-c)[p_x(y+b) - p_x(y-b)]}, \\ &\qquad c \leq x \leq d \\ &\qquad x - b \leq y \leq x + b \\ &= 0, \text{ otherwise} \end{aligned} \tag{A.8}$$

A representative density function of this nature is shown in Figure 20. Note that for any particular measurement, say y = k, the

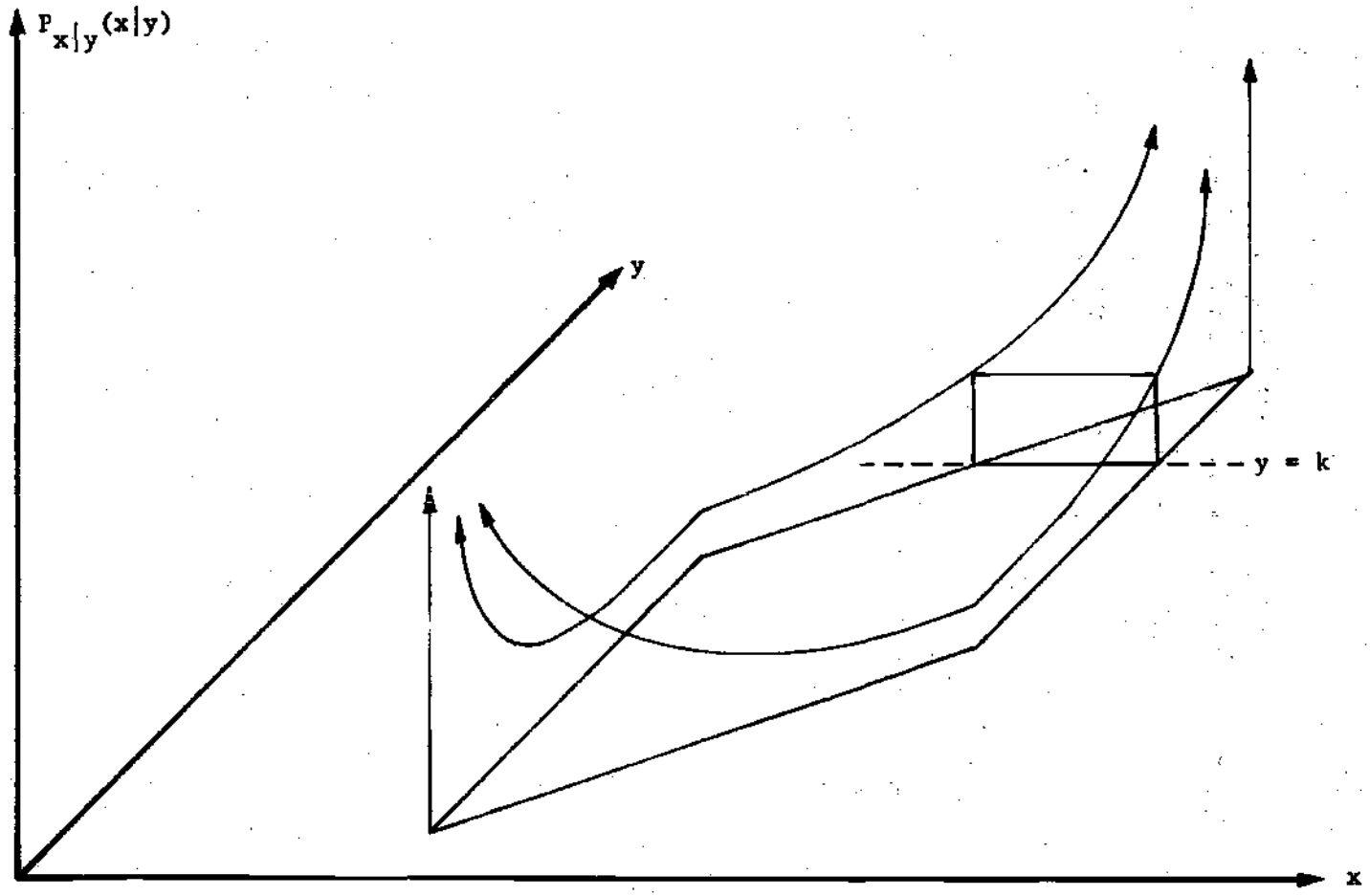



Figure 20. Representative Conditional Density

resulting conditional density for x is uniform and bounded and therefore is symmetrical. Also note that for a particular measurement, say y = k, the nonzero region of definition for $p_{x|y}(x|k)$ is the intersection of reachable states and states which could have given the particular measurement.

# BIBLIOGRAPHY


1. Norbert Weiner: The Extrapolation, Interpolation, and Smoothing of Stationary Time Series, John Wiley and Sons, New York, 1949.

2. R. E. Kalman: "A New Approach to Linear Filtering and Control Problems," ASME Transactions, Series D, Journal of Basic Engineering, Vol. 82, March, 1960.

3. R. C. Booton: "Final-Value Systems with Gaussian Inputs," IRE Transactions on Information Theory, Vol. IT-2, 1956.

4. Leonard G. Shaw: "Optimum Stochastic Control," Disciplines and Techniques of Systems Control, Chapter 4, Blaisdell Publishing Company, New York, 1965.

5. David Middleton: Introduction to Statistical Communication Theory, McGraw-Hill Book Company, New York, 1960.

6. Samuel L. Fagin: "Recursive Linear Regression Theory, Optimal Filter Theory, and Error Analysis of Optimal System Theory," IEEE International Convention Record, Part I, 1964.

7. S. F. Schmidt: "State Space Techniques Applied to the Design of a Space Navigation System," JACC Preprints, 1962.

8. G. L. Park: "A Sample Average Filter for Linear Control Systems," JACC, Troy, New York, 1965.

9. J. Heffes: "The Effect of Erroneous Models on the Kalman Filter Response," IEEE Transactions on Automatic Control, Vol. AC-4, July, 1966.

10. T. Nishimura: "Error Bounds of Continuous Kalman Filters and the Application to Orbit Determination Problems," IEEE Transactions on Automatic Control, Vol. AC-12, June, 1967.

11. R. J. Fitzgerald: "Divergence of the Kalman Filter," IEEE Transactions on Automatic Control, Vol. AC-16, December, 1971.

12. Robert C. K. Lee: Optimal Estimation, Identification, and Control, Research Monograph No. 28, M.I.T. Press, Cambridge, Massachusetts, 1964.

13. H. Cox: "On the Estimation of State Variables and Parameters for Noisy Dynamic Systems," IEEE Transactions on Automatic Control, Vol. AC-9, 1964.

14. M. Athans, R. P. Wishner, and A. Bertolini: "Suboptimal State Estimation for Continuous-Time Nonlinear Systems from Discrete Noisy Measurements," IEEE Transactions on Automatic Control, Vol. AC-13, October, 1968.

15. Paul A. Frost and Thomas Kailath: "An Innovations Approach to Least Squares Estimation - Part III: Nonlinear Estimation in White Gaussian Noise," IEEE Transactions on Automatic Control, Vol. AC-16, June, 1971.

16. Raman K. Mehra: "A Comparison of Several Nonlinear Filters for Reentry Vehicle Tracking," IEEE Transactions on Automatic Control, Vol. AC-16, August, 1971.

17. Te-Son Kuo: "A Moment Technique for Suboptimal Adaptive Nonlinear Filtering," Ph.D. Thesis, Georgia Institute of Technology, December, 1969.

18. F. C. Schweppe: "Recursive State Estimation: Unknown but Bounded Errors and System Inputs," IEEE Transactions on Automatic Control, Vol. AC-13, February, 1968.

19. A. P. Sage: Optimum Systems Control, Prentice-Hall, 1968.

20. A. Papoulis: Probability, Random Variables, and Stochastic Processes, McGraw-Hill Book Company, New York, 1965.

21. Y. Sunahara: "An Approximate Method of State Estimation for Nonlinear Dynamical Systems," Preprints of the 1969 Joint Automatic Control Conference, Boulder, Colorado.

22. Te-Son Kuo and J. R. Rowland: "A Moment Technique for Suboptimal Adaptive Nonlinear Filtering," 1970 Southwestern I.E.E.E. Conference, April, 1970.

23. J. D. Glover and F. C. Schweppe: "Control of Linear Dynamic Systems with Set Constrained Disturbances," IEEE Transactions on Automatic Control, Vol. AC-16, October, 1971.

24. George M. Clark: "An Improved Algorithm for the Combined Estimation and Control of Nongaussian Stochastic Systems," Ph.D. Thesis, Georgia Institute of Technology, May, 1971.

25. Robert G. Gallagher: Information Theory and Reliable Communication, John Wiley and Sons, New York, 1968.

# VITA

Richard Erwin Bryan was born in Asheville, North Carolina, on April 19, 1939. He attended public schools in Asheville and graduated from high school in 1957. In the same year, he entered the Georgia Institute of Technology and received the B.E.E. degree in June, 1962. During this period at Georgia Tech, he worked as a co-op student at General Dynamics/Fort Worth.

Mr. Bryan entered graduate school at Georgia Tech in September, 1962, and was employed concurrently as a research assistant in charge of the Analog Computation Laboratory at the Georgia Tech Engineering Experiment Station. He received the M.S.E.E. degree in June, 1965. From September, 1965, to September, 1966, he was employed as a project director in charge of Naval Mine Defense research at the Georgia Tech Engineering Experiment Station. From September, 1966, to September, 1969, he held the position of Instructor in the Georgia Tech School of Electrical Engineering while working toward his doctorate. In September, 1969, he assumed the position of Assistant Professor of Quantitative Methods at Georgia State University and received the Ph.D. degree in June, 1972.

Mr. Bryan is married to the former Beverly Branson of Columbia, South Carolina. They were married February 14, 1970.